DOCUMENT RESUME

ED 063 009 LI 003 692

AUTHOR	Johnson, Donald W.
TITLE	Toward a National Serials Data Program: Final Report of the National Serials Pilot Project.
INSTITUTION	Association of Research Libraries, Washington, D.C.
PUB DATE	72
NOTE	94p.; 4 References
AVAILABLE FROM	The Association of Research Libraries, 1527 New Hampshire Avenue, N.W., Washington, D.C. 20036 ($5.00 prepaid)
EDRS PRICE	MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS	*Automation; *Data Bases; National Libraries; Periodicals; *Serials; Systems Development
IDENTIFIERS	*Machine Readable Bibliographic Data Bases; MARC

ABSTRACT

The National Serials Pilot Project, Phase II of the National Serials Data Program, is described. Utilizing the MARC format for processing serials, the objectives were: (1) to create a machine-readable file containing live serials in the fields of science and technology, (2) to produce a number of preliminary listings, and (3) to produce one or more written reports covering procedures, problems and results. Data were input via an administrative terminal system to a 360/40 computer; processing of data was done on a 360/50 computer. Among the conclusions and recommendations are: (1) a national serials data bank in machine-readable form is both technically and economically feasible; (2) such a data bank should have its own machine-readable authority file for corporate names; (3) input and output in upper case only would be more satisfactory from both the systems viewpoint and the cost viewpoint, but probably would not be accepted by the library community; and (4) serious consideration should be given to the question of applicability of existing cataloging rules in the determination of main entry in a machine-readable file. (Author)

ED 063009

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY
Association of Research Libraries
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

TOWARD A NATIONAL SERIALS DATA PROGRAM:

Final Report of the National Serials Pilot Project

By

Donald W. Johnson

The Association of Research Libraries
1527 New Hampshire Avenue, N.W.
Washington, D.C.
1972

LI 003 692

TABLE OF CONTENTS

Page

PREFACE vii

ACKNOWLEDGEMENTS ix

ABBREVIATIONS xi

ABSTRACT xiii

INTRODUCTION 1

METHODS AND PROCEDURES 5

- National Libraries Task Force 5
- Staff 5
- System 7
- Character Representation 7
- Selection of Titles 8
- File Building 9
- Authorities 10
- Innovations in MARC 11
- Standard Serial Numbers 12
- Programming and Program Maintenance 12

Page

RESULTS 15

Objectives 15

Materials Delivered to National Libraries 15

Systems Documentation 16

Master File Data 20

CONCLUSIONS 21

Feasibility 21

Authority File 21

Rules of Entry 22

Cost of Combined Upper and Lower Case Output 22

Expectations 23

RECOMMENDATIONS 25

National Serials Data Program 25

Systems 25

MARC Serials Format 25

Program Improvements 26

Authority File 26

Input and Output 26

Rules of Entry 27

MARC Serials Format Improvements 27

BIBLIOGRAPHY 29

APPENDIX A: SUGGESTIONS FOR FUTURE ORGANIZATION AND ADMINISTRATION OF A NATIONAL SERIALS DATA PROGRAM 31

APPENDIX B: SUGGESTED IMPROVEMENTS IN THE MARC SERIALS FORMAT 37

Page

APPENDIX C: SUMMARY OF NATIONAL SERIALS DATA PROGRAM 41

APPENDIX D: OPERATIONS FLOWCHARTS 47

CHARTS

Chart I: National Serials Pilot Project Administrative Organization Chart 6

Chart II: National Serials Data Program Proposed Administrative Organization Chart 35

TABLES

Table I: General File Characteristics 86

Table II: Frequency Distribution of Variable Fields in File 87

Table III: Languages of Publication 89-

Table IV: Countries of Publication 90

PREFACE

The Association of Research Libraries is pleased to issue this report on the National Serials Pilot Project, which the Association administered on behalf of the three national libraries and at their request. Active work on the pilot project extended over a period of a little more than a year and a half and ended in June 1971. It was the judgment of those directly involved that the pilot phase of the National Serials Data Program had been completed by that time.

Although the Association of Research Libraries had no special capability to prepare a machine-readable record of serials, it did agree to undertake administrative responsibility for an experimental period as a means of furthering work towards the development of a national serials record. Because of its experimental nature, it was considered best to have as manager of the project an agency not directly involved. Policy guidance, however, was provided by the three national libraries through their representatives on the National Libraries' Task Force.

This report was prepared by the director of the pilot project, Donald W. Johnson, as a record of the effort to create a serials record in machine-readable form, and of the experiences, problems and findings that resulted from that effort. The task was undertaken as an actual operating experience from which, it was hoped, much could be learned about problems to be solved, procedures and methods to be followed, and guidelines to be developed for the main undertaking of a national serials record. The pilot project served this purpose.

After the work was completed, copies of all the materials generated by the project were deposited with the three national libraries and with the Association of Research Libraries. They include a copy of the machine-readable tape representing the file of data developed, the full range of computer programs used, and an extended manual of procedures to serve as a guide for the National Serials Data Program.

Appendix A, Summary of National Serials Data Program, is not strictly speaking a part of the report. This statement of policies for the National Serials Data Program was prepared by the Library of Congress, with the concurrence of the National Agricultural Library and the National Library of Medicine, as a report to the library community on the next steps that will be undertaken to develop a national serials record. It is based on the findings of the pilot project. Thus, although not itself an integral part of the report, it is a significant addition to it since it indicates the program that will be followed by the three national libraries in the years ahead in creating the serials record. The Association of Research Libraries is indebted to the Library of Congress for the preparation of this statement and for permission to include it as a part of this report.

It is now assured that a national serials data record will be developed in the light of the best thinking of experts and of the findings of the experimental project. Such a record will benefit library users in this country and abroad. The Association of Research Libraries is pleased to have had the opportunity to make a modest contribution towards the development of this record.

The Association wishes to take this occasion to express its thanks to the three national libraries and the Council on Library Resources for their confidence in the ARL and their assistance in making the National Serials Pilot Project serve the experimental purposes for which it was intended.

The recommendations of the report are the responsibility of the director of the project and the statement of plans for future development is the responsibility of the three national libraries. They do not necessarily reflect the views of the ARL.

Stephen A. McCarthy
Executive Director
Association of Research Libraries
March 23, 1972

ACKNOWLEDGEMENTS

Many individuals have contributed to the success of the National Serials Pilot Project. They include the members of the National Libraries' Task Force, the ARL advisory committee to the project, staff members of the three national libraries, ARL administrative personnel, and many others. The members of the ARL Advisory Committee to the National Serials Pilot Project were: Mr. William Budington of the John Crerar Library (Chairman); Mr. Roy Kidman of Rutgers and, more recently, the University of Southern California; Mr. John McGowan of Northwestern University; and Mr. Jerrold Orne of the University of North Carolina.

Two people from the Council on Library Resources, by virtue of the number and importance of their contributions, must be mentioned: Mr. Lawrence Livingston and Mr. George Parsons. Suffice it to say that the NSPP probably would have foundered but for their efforts and good offices.

Among the members of the NSPP staff whose contributions have been most significant are Mr. Richard L. Andrews, programmer; Mrs. Tillie Krieger, research associate; and Mrs. Sigrid Slivka, research associate. The bulk of Mr. Andrews' contributions will be found in the systems documentation. Mrs. Krieger helped in the early stages of development of the manual of procedures, while Mrs. Slivka is largely responsible for the operations flowcharts contained in that manual. (References to the manual and system documentation will be found in the Bibliography.)

To all of these, and to many others who have not been named, go the sincere and warm thanks of the project director.

Sole responsibility for any shortcomings of either the project or the reports emanating from it must rest on his shoulders.

ABBREVIATIONS

ALA	American Library Association
ARL	Association of Research Libraries
ATS	Administrative Terminal System
CAIN	Cataloging and Indexing
CLR	Council on Library Resources
COBOL	Common Business Oriented Language
COSATI	Committee on Scientific and Technical Information
DOS	Disk Operating System
EDP	Electronic Data Processing
ISO	Information Systems Office (Library of Congress)
JCULS	Joint Committee on the Union List of Serials
LC	Library of Congress
MARC	Machine Readable Cataloging
MED	Mechanized Error Detection
NAL	National Agricultural Library
NLM	National Library of Medicine
NLTF	U.S. National Libraries Task Force on Automation and Other Cooperative Services
NSF	National Science Foundation
NSPP	National Serials Pilot Project
OS	Operating System
QCB	Quality Control Block
SSN	Standard Serial Number
UCMP	Union Catalog of Medical Periodicals

ABSTRACT

The National Serials Pilot Project, Phase II of the National Serials Data Program, is described. The project was funded by the National Agricultural Library, the Council on Library Resources, the Library of Congress, and the National Library of Medicine. It was administered by the Association of Research Libraries under policy direction from the U. S. National Libraries' Task Force on Automation and Other Cooperative Services.

Utilizing the MARC format for processing serials, the objectives were 1) to create a machine-readable file containing live serials in the fields of science and technology, 2) to produce a number of preliminary listings, and 3) to produce one or more written reports covering procedures, problems and results. Data were input via an administrative terminal system to a 360/40 computer; processing of data was done on a 360/50 computer.

Among the conclusions and recommendations are: 1) a national serials data bank in machine-readable form is both technically and economically feasible; 2) such a data bank should have its own machine-readable authority file for corporate names; 3) input and output in upper case only would be more satisfactory from both the systems viewpoint and the cost viewpoint, but probably would not be accepted by the library community; and 4) serious consideration should be given to the question of applicability of existing cataloging rules in the determination of "main entry" in a machine-readable file.

INTRODUCTION

Serials have long constituted a special problem for libraries, a problem very much aggravated by the "information explosion" following World War II. Recognizing this fact, in April 1964 the Committee on Scientific and Technical Information (COSATI) created a special task force to study ways to improve the processing and utilization of journal literature. Its work prompted the National Science Foundation to award a study contract to the Information Dynamics Corporation in April 1965. The study report proposed a computer-based serials data program for science and technology (1).

Discussions ensued involving the three national libraries, the NSF, and the Association of Research Libraries (ARL). In June 1966, the ARL appointed an ad hoc committee to explore the matter; this ad hoc committee subsequently became the Subcommittee on World List of Serials of the Joint Committee on the Union List of Serials (JCULS). In December 1966, the subcommittee asked the Library of Congress to prepare a proposal to be considered during the 1967 midwinter meeting of the American Library Association. Later that year, at the annual conference of the ALA, an announcement was made that the directors of the three national libraries had agreed to undertake a cooperative serials data program. In August of that year a working paper was completed, and by October funds for what was to be Phase I were obtained.

Phase I began in January 1968 with the Library of Congress, through its Information Systems Office, as executive agent. The JCULS acted in an advisory capacity for the work on Phase I. In this phase the objectives were to define serials, identify the data elements needed to control them, and develop a content format for serials. This last resulted in the MARC serials format. A user survey was also undertaken as part of Phase I. The final report of Phase I was published in early 1969 (2). It recommended that a pilot project should be begun as soon as possible.

In May 1969 the U.S. National Libraries' Task Force on Automation and Other Cooperative Services (NLTF) was authorized by the directors of the three national libraries to seek funds for a serials pilot project. The National Agricultural Library (NAL) agreed to provide $100,007 to get such a project started. The directors of the three national libraries then approached the ARL and, after some disucssions it was agreed that the ARL would administer the project on behalf of the three national libraries, under policy direction of the NLTF.

Unfortunately, the original agreement did not include a definition of "policy" as interpreted for this project, nor did it provide for planning and funding of subsequent phases of a national serials data program. In addition to creating an undesirable working atmosphere for the staff, these omissions made it necessary to devote a substantial amount of staff time to the preparation of proposals for the continuation of the project.

By December 1970 the funds provided by the NAL had been exhausted. Anticipating this need, the three national libraries and the Council on Library Resources had agreed in November 1970 that each of the three libraries would provide $2,000 per month for the seven months remaining to the project and that the CLR would contribute an amount not to exceed $19,107. Total funding for the project, then, was $161,114, expended over a span of 21.5 months.

By September 1969 the ARL had employed a director for the project. Among his first tasks was to meet with the NLTF and to arrive at an agreement upon realizable goals. These were defined on October 16, 1969, as: 1) to create a machine-readable file containing live serials in the fields of science and technology held by the three national libraries; 2) to produce a number of preliminary listings, including a union list and other lists of interest to management; and 3) to produce one or more written reports relating to problems, solutions, information regarding the universe of serials, and recommendations. A serial was defined as "a publication in successive parts bearing numerical or chronological designations and intended to be continued indefinitely."

The NLTF designated Mr. Lawrence Livingston of the Council on Library Resources as liaison between the NLTF and the project director.

With these objectives in mind, the director then developed a proposed system for the preparation of a serial record in machine-readable form and presented it to the NLTF for approval. Essentially, the system employed fixed-variable fields, with provision for carrying all data peculiar to the MARC serials format in the records so that conversion to a MARC file could be accomplished with a minimum of time, cost, and difficulty. The proposed system, however, was not itself a MARC-formatted system. After extended discussion, the NLTF approved the system proposed by the director, but this decision was later rescinded.

In a subsequent review of this decision by the advisory committee to the project, it was pointed out that the decision of the Task Force had been in large part based upon incomplete information regarding available MARC programs and their capabilities, and the difficulty of securing competent programming support. Later, at a joint meeting of the advisory committee and the NLTF, the chairman of the NLTF announced that programming assistance from the LC would be provided. This announcement led to the re-opening of the question of the system to be used. It was agreed that if the necessary programming assistance were available, it would be preferable to use the MARC format.

On December 5, 1969, a meeting was held in the office of the Deputy Librarian of Congress at which it was agreed that the LC would provide MARC internal format programs rewritten in COBOL, together with disk storage for an administrative terminal system, and the National Library of Medicine would make its computer available. Preliminary versions of the COBOL programs had been completed and were being de-bugged on the NLM computer. The National Serials Pilot Project thus provided a field test for the MARC serials format in addition to its other tasks.

METHODS AND PROCEDURES

National Libraries Task Force

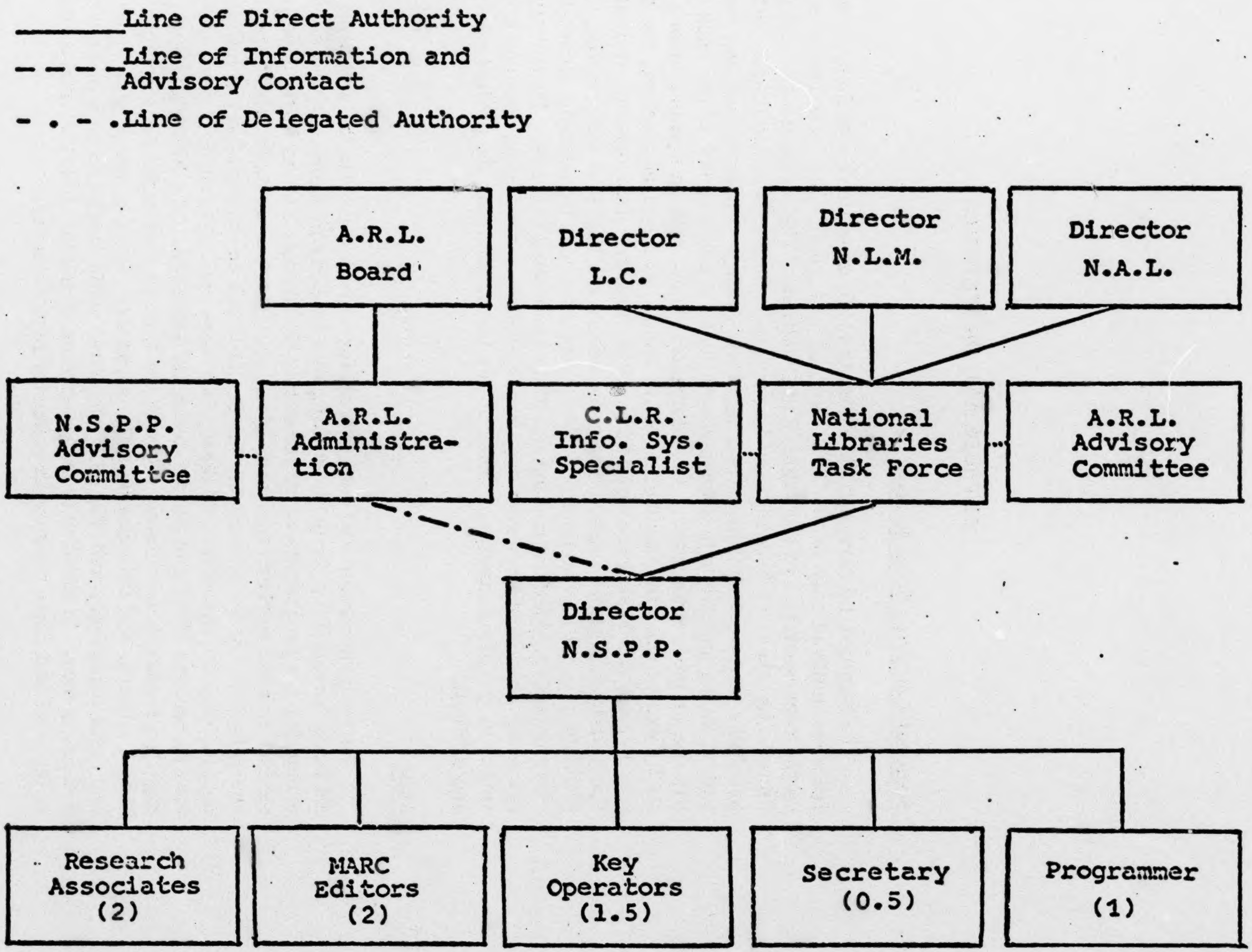
The National Libraries Task Force is made up of three members, one chosen from the staffs of each of the three national libraries. The Task Force is chaired by the representative of the Library of Congress. The members are selected for their comprehensive knowledge of library operations, their familiarity with the problems of the national libraries, and their ability to analyze common problems and to devise and evaluate proposed solutions. Each member of the Task Force has one or more alternates. The alternates do not vote on matters before the NLTF except when the member is not present. At a typical meeting, the three members of the Task Force will be present, together with several alternates, specialists in matters expected to come before the NLTF at that meeting, and possibly several invited outsiders whose opinions may be wanted. Under such circumstances, decision making is a time consuming process. Frequently, decisions represent compromises, especially when conflicting opinions are presented by qualified experts. Thus, policy formulation for the project was slow and difficult.

Staff

Since librarianship and data processing were involved in the activities of the National Serials Pilot Project, staffing was an early problem. It was solved, except for the programmer, by recruiting librarians, library assistants and skilled clerical workers who were given on-the-job training for their specific assignments. At the maximum, the project staff consisted of seven full-time people, plus the director. Of these, two were professional librarians whose experience was chiefly in library technical services. The librarians were designated research associates. Two of the library assistants were designated MARC editors, and two became ATS operators. Finally, the programmer had had prior experience with the NLM computer while employed on a project of a federal agency. In addition to the programmer and the director, three members of the staff had taken courses in computer programming.

CHART I

NATIONAL SERIALS PILOT PROJECT
ADMINISTRATIVE ORGANIZATION CHART


Line of Direct Authority
Line of Information and Advisory Contact
Line of Delegated Authority
A.R.L. Board
Director L.C.
Director N.L.M.
Director N.A.L.
N.S.P.P. Advisory Committee
A.R.L. Administration
C.L.R. Info. Sys. Specialist
National Libraries Task Force
A.R.L. Advisory Committee
Director N.S.P.P.
Research Associates (2)
MARC Editors (2)
Key Operators (1.5)
Secretary (0.5)
Programmer (1)

The research associates, although skilled in bibliographic verification, required specific training in the requirements of a MARC oriented system as well as general training in computer concepts, terminology, etc. The MARC editors were given comprehensive training in the application of the MARC serials format; this was made somewhat easier by the fact that both already had some understanding of library science. Since experienced ATS operators were not available, the staff had to be trained, and the programmer had to be given a condensed and selective course in library operations.

Training for specific assignments was supplemented by explanations of the inter-relationships among the several tasks involved to assure effective performance. The recipe was to take existing skills, add other skills, mix with an understanding of the goals to be achieved, and produce a group that could and would function as a team.

System

The National Serials Pilot Project employed the MARC serials processing (or internal) format. The full format was used except for those fields intended for representation of subject headings or for printing of catalog cards. Input was via ATS to a 360/40 DOS (Disk Operating System) computer at the LC. Typewriter terminals using acoustic couplers and telephone lines linked to data sets at the LC were employed. For input purposes, the NSPP was in an on-line mode with the LC computer.

Processing of data was done on an IBM 360/50 OS (Operating System) computer at the NLM. Generally, processing was done once a week from a magnetic tape prepared at the LC from data input during that period by the NSPP. This tape was transferred to the NLM, where the systems programs and files were maintained. Thus, for processing purposes, the NSPP was in batch mode.

Character Representation

The hardware configuration used imposes constraints on any computerized record system. In the case of the pilot project, this condition was emphasized by the need to employ two computers with differeing capabilities. For example, the LC computer had an expanded character set that would have permitted printing almost all modified letters in the Roman alphabet; the NLM computer lacked this capability. Since diagnostic listings showing precisely what had been input were to be produced by the NLM computer, the limitations of this computer were determining. Thus, all character modifications were omitted. This omission was considered preferable to the development of a code system for modified letters because of the numerous and costly program changes its use would have entailed.

Whether to input in upper case only or in upper and lower case had to be decided at a very early stage. There is a natural preference for both upper and lower case. Apart from this fact, however, electing to use upper and lower case preserved the option of being able to output either in upper case or in upper and lower case, while opting for upper-case input would have proscribed any such subsequent option. It was decided, therefore, that input would be in both upper and lower case.

Selection of Titles

Titles had somehow to be selected for inclusion in the file. Several alternatives were available. Since the NSPP was working with the three national libraries and would be concerned with their holdings, basing the file on their serial records seemed to be plausible. The chief objections to this approach were: 1) the scientific and technical serials with which the NSPP was to concern itself would first have to be identified in those files, i. e., a list of titles to be included would first have to be prepared; and 2) the serials files were in constant use by many people and addition of NSPP staff to the users of these files would have created problems for both the libraries and the project staff.

A second alternative was to base the file on an already established, published list. This had some obvious merit, but consideration of it led quickly to the third alternative: the use of an established list that was also available in machine-readable form which, by reformatting, would reduce the magnitude of the task. The fact that the National Science Library of Canada had already offered tapes of its *Union List of Scientific Serials in Canadian Libraries*, third edition (acronym: *Canser 3*) made this even more appealing, particularly since its subject limitations corresponded to those of the NSPP. It was decided, therefore, to accept the offer of the National Science Library of Canada and to reformat its file to ATS output format, relying upon software to convert the resulting file into a MARC format.

This was a sound decision based upon then-known facts and expectations. That it did not work very well in practice is attributable to a number of factors. First of all, at the time the decision was made it was expected that the NSPP would be able to be in an on-line mode for at least portions of the reformatted file. If this had proved feasible, key input by the NSPP would have been substantially reduced and the reformatting would have been worthwhile. But it never became possible to read blocks of records from the reformatted file into the LC computer so that a reformatted *Canser 3* file could be accessed via ATS terminals. Secondly, it developed that there was not really any truly accurate and complete description of the characteristics of *Canser 3* available. The file had originally been built in the mid 1950s and, as is so often the case, no comprehensive documentation existed. In any event, there had been many changes since then, including changes in hardware and in input methods. Starting with the information supplied, reformatting was begun. When reformatting based upon that information failed to yield the desired results, a great deal of money and time had already been invested and there was a natural reluctance to begin from scratch with some other list, possibly only to encounter the same kinds of difficulties. Eventually, the task was accomplished, but only after much groping and the passage of seven months.

Knowing what is now known, a reformatting approach probably would not have been chosen, especially since MARC was involved. Since the data eventually had to be keyed, the reformatting really did little good. From this experience the conviction has grown that it is not possible to use reformatting alone to convert from any non-MARC format to MARC owing to the details required by MARC. In other words, since MARC requires fixed-field coding, subfield coding, primary and secondary indicators, and various other things not contained in a non-MARC file, identification of needed data cannot be complete; hence, human intervention is required to complete the conversion to MARC. This being the case, the reformatting approach for NSPP purposes must be regarded as suspect and, in this specific instance, a costly error. This is not, however, to be construed as a rejection of reformatting _per se_.

In addition to the _Canser 3_ titles to be input to the file, it was decided to include _Index Medicus_ titles that had been reported to the _Union Catalog of Medical Periodicals_ and for which data were available on magnetic tape in the UCMP format. These titles had also to be reformatted to the ATS output format, but this was a much simpler job and was quickly accomplished.

Finally, the Cataloging and Indexing file maintained in machine-readable form by the NAL was reformatted by its staff to the ATS output format and these titles were also included in the file.

File Building

Because of the delays encountered in reformatting _Canser 3_, file building began with _Index Medicus_ titles not contained in _Canser 3_. It was expected that the others would be incorporated later. All of the data for each title had to be searched and verified, MARC-edited, keyed, printed with line numbers via ATS, revised, corrected, and sent to queue before being processed into the master file. (For operations flowcharts, see Appendix D.) These activities were made more difficult in the early stages because the package of programs delivered in January 1970 had not been fully tested and de-bugged, and the LC programmers were continuing to make corrections in it. On numerous occasions program changes accidentally destroyed portions of the file or portions of records in the file. Consequently, progress until the end of 1970 was quite slow.

MARC-editing of serials proved to be far more complex and time consuming than MARC-editing of monographs. The difficulties were attributable to the nature of serials and the NSPP practice of building a separate record for each title, including all titles in the bibliographic trail of an entity. This type of record was considered preferable in a national serials data bank to the "consolidated entry." Thus the editing process required identification of all links in a bibliographic chain and the editing of these elements to represent all significant relationships. This proved to be a complex task.

Authorities

All of the non-Canser 3, Index Medicus titles had been input to the file when it was decided by the NLTF that the LC authority files were to be the primary (but not the sole) authority for all records, so far as choice of entry and form of name were concerned. This decision necessitated the scrapping of the existing file and beginning anew, since none of these titles had been searched at the LC because all of them could be found at the NLM.

By the time the pilot project had been completed, a total of thirty three different authorities had been used for records (or portions of records) in the file. At no time did the project have its own authority file. Inevitably, there were conflicts, and these were not only between authorities but within them as well.

One of the principal reasons for the decision to use the files of the LC as primary authority was the hope that it would be possible to build a consistent data base if authorities were limited. This might have been realizable if it had been possible to limit searching and verification to a single authority file and if authority files were all that they ideally should be. But many titles were not to be found in the authority files of the LC, and it is doubtful that any authority file anywhere is entirely consistent. The problem was aggravated by the practice known as "superimposition."

One of the differences between the Anglo-American Cataloging Rules and previous cataloging practice relates to entry under place. The change in this respect created problems for all libraries, but especially for the LC. Since it was not feasible for LC to make all the necessary changes in its authority files, entries for newly established corporate headings are made in accordance with the new rules, while headings previously established according to the old rules coexist in the same files. There is the additional problem of distinguishing the "superimposed" headings in the authority files from the others.

Under these conditions, the results were bound to be inconsistent, and the decision to regard the authority files of the LC as primary authority was rescinded by the NLTF on April 6, 1971. At the same time, the NLTF decided that the data supplied by the first of the three national libraries to report any given title were to be accepted as authority. Although this reduced the time invested in bibliographic verification, it did not make the NSPP file more consistent.

Innovations in MARC

Modifications in a complex machine-readable bibliographic format such as MARC, which has been designed as a standard, cannot be made without serious consequences. If changes or substitutions are made in MARC by a library, its file is likely to be incompatible with MARC files in other libraries. Adoption of this practice by libraries militates against the development and acceptance of a standard and thus defeats the effort to achieve a national bibliographic data bank in machine-readable form. Lack of standardization prevents or inhibits the ready interchange of information, the addition of information from various sources, and the ability to respond promptly and accurately to inquiries.

In keeping with the objectives of the NSPP, no alterations were made in the basic MARC format or in the meaning of its tags, subfields or indicators, but the following additions were made, for the reasons given:

1) MARC records carry a machine-provided date-entered on file. This item is changed each time any change is made in a record. Since it could be important to have such dates on a field-by-field basis in a serials file, a subfield was added to each variable field for this purpose. This additional subfield contains the date on which that particular data element was last entered or updated.

2) A second subfield for each variable field was created to make it possible to trace a data element back to the authority that had been accepted for that element and its form. In this subfield, a numeric code was entered that indicated which of thirty three sources had been accepted as authority for that data element.

3) In the recording of holdings, the customary methods seemed inappropriate. For one thing, as file size grew ever larger it could be expected that file space required for holdings data would become prohibitive even if only a few libraries were to be represented. For another thing, conventional methods of representing holdings data in union lists would imply a need for continuous updating, requiring permanent, formal reporting of all changes

in holdings by each participating library. Such an arrangement, even if it could be established, would not be maintained nor would it be equally honored by all participating libraries. Therefore, a simple system for encoding holdings data for each library was developed. A subfield was added to field 850, and in this subfield a "0" was entered if the extent of a library's holdings was not known; a "1" if its holdings were complete or substantially complete; a "2" if its holdings were substantially incomplete; a "3" if a title was held for a limited time only; or a "4" if a title was received by the library but not held at all. (This was developed as an experiment, in addition to the conventional holdings record. It is understood that this system has been adopted by others.)

4) To each of the linking-entry fields a subfield was added to which was input, as appropriate and as data were available, a numeric link rather than alphabetical data. The number contained in this subfield corresponded to the number carried in field 035 (local system number) of the record to which reference was being made. Since this number contained a modulus-11 check digit, it was machine-checkable. It had the further advantage that it conserved file space while eliminating failures to match because of errors and/or discrepancies in the alphabetical representations of the two records. The use of numbers in this way follows logic inherent in the MARC serials format, and thus was really an extension of the format rather than an addition to it.

Standard Serial Numbers

In January 1970, it was decided by the NLTF and the ARL advisory committee that the NSPP should experiment with standard serial numbers. Since, however, the NSPP could have no authority for permanent assignment of SSNs, they were assigned as "local system numbers," employing field 035. Programs were developed to produce a list of valid SSNs, beginning with any given number and ending with any given number, and to check for the validity, omission, or duplication of SSNs in the file.

Programming and Program Maintenance

The package of programs, rewritten for the NSPP use in COBOL F and officially given to the project in January 1970, was not operational until mid-April. Even then, there were many serious errors in the programs and in their logic and, as has been noted elsewhere, efforts to correct these errors had on several occasions impaired the operations and records of the project.

The need for a full-time programmer on the project staff, partly for program correction and maintenance and partly to develop other programs, became evident in the course of the summer and a programmer was appointed in September. Two of his more important contributions were: 1) the development of a mechanized error detection (MED) suite of programs which tested each record in the master file for each of 265 machine-detectable errors and produced error messages as appropriate; and 2) documentation of all systems programs used by the project with the aid of Autoflow, thus producing program-logic flowcharts. These have been produced as a separate report (3).

It should be mentioned that both programming and program maintenance would have been simpler had the NSPP been in on-line mode with the computer used for processing its data. In that event, one of the more important objectives of an administrative terminal system could have been achieved, namely inputting and de-bugging programs and/or program changes via remote access terminal and being able to do so during normal working hours.

RESULTS

Objectives

At the conclusion of the pilot project, all the necessary systems work had been completed; all programs needed for achieving the stated objectives had been written; a file containing 7,049 records had been built; and listings and reports had been produced. One of the more important of these reports was the Manual of Procedures (4).

The tasks defined for the NSPP by the NLTF on October 16, 1969, had all been achieved, although the union list mentioned at the time was never produced, since holdings data were not reported to the project. Such holdings information as is contained in the master file is incomplete and does not reflect the holdings of each of the three national libraries of the titles in the master file. These holdings records can be added by the three national libraries at any time. The program capabilities already exist for producing a union list.

Materials Delivered to National Libraries

Each of the three national libraries was provided with the following:

1) magnetic tapes:
    a) NSPP master file,
    b) Canser 3 (reformatted),
    c) Canser 3 (unreformatted),
    d) Index Medicus titles from UCMP (reformatted),
    e) CAIN (reformatted), and
    f) NLM non-Index Medicus titles from UCMP (reformatted);

2) printouts:

  a) NSPP master file,
  b) program listings,
  c) _Canser_ _3_ (reformatted),
  d) _Index_ _Medicus_ titles from UCMP (reformatted),
  e) CAIN (reformatted),
  f) NLM non-_Index_ _Medicus_ titles from UCMP (reformatted), and
  g) checklist based upon NSPP master file; and

3) documentation:

  a) _Manual_ _of_ _Procedures_,
  b) systems documentation in flowchart form,
  c) program source decks, and
  d) final report.

The master file presented to each of the three national libraries, both on tape and in printed form, was first tested for more than two million machine-detectable errors, with negative results, i.e., none of these errors was present in the master file.

The master file, reformatted _Canser_ _3_, reformatted _Index_ _Medicus_ by way of UCMP, reformatted non-_Index_ _Medicus_ by way of UCMP, and CAIN files together comprise more than 40,000 scientific and technical titles. All of these were presented either in the MARC format or in the ATS output format. They are readily convertible to MARC with the software presented.

Systems Documentation

For each program included in the systems documentation, the following information was given.

1) general description, including data sets used, inputs and outputs;
2) program listing in COBOL F;

3) procedural statement label index, alphabetically arranged;
4) tables of contents and references;
5) program logic flowchart of the procedure division;
6) procedure division analysis;
7) data cross references, alphabetically arranged by data referred to;
8) data record map; and
9) data division index.

The following programs were included in the systems documentation.

1) BIBSYS. This is a COBOL program which reads a tape of ATS records (133 characters each), edits each record for content, and builds an output tape which consists of fixed-length records in MARC format. The maximum output record length is 4,088 bytes. The program is generalized to allow for fluctuations in user formats. Through a series of control cards which are inputs to the program, the allowable fixed fields and allowable tags are defined. The input will be edited accordingly. As many control cards as are necessary to define the tags and fixed fields are permitted.

2) BIBSYS 1. This COBOL program updates the master data base. The input to the program is the transaction file produced by BIBSYS which has been sorted by the record number. One control card is also an input to the program; this card specifies whether there is no old master data base or whether this is a normal update run against an old master. The master records are updated on the record level, tag level, or subfield level, and a new master data base is produced.

3) BIBSHOW. This is a COBOL program that produces a diagnostic-history listing of the maintenance run. Inputs to the program are:

   a) the summary record file produced by the BIBSYS 1 program, and

   b) the same parameter cards used for BIBSYS processing.

The program accepts at the logical record level each transaction introduced into the maintenance run, translating into readable text a message indicating whether valid or invalid processing has occurred. If a change has been made in any portion of a master record, the program will display, in tag variable field form, the final appearance of the updated master record. The parameter cards are used by the program to translate the internal record ingredients back to the original tag name and fixed-field definitions required for error-detection processing.

4) BIBSKED. This is a generalized key-building routine that takes the tagged information (as per control cards) and builds a sort key that is appended to the MARC record data in each record. Up to six sort keys may be specified, either fixed fields or variable fields, or both. Either all records may be stipulated or, if a "pull" card is included with the parameter cards, one field may be stipulated as a criterion for building a miniaturized data base with sort key appended.

5) BIBSTRP. This program rebuilds a MARC record from the data set of 2,522 character records passed to it from BIBSKED and the sort. In building this record, it discards both the appended sort key and the quality control block (QCB) from the input.

6) GENPRNTA. This is a two-program utility package able to perform a variety of computer printing tasks on an automated file of bibliographic information carried in the MARC format. It accepts the MARC file from magnetic tape and report parameter cards from the system's card reader, matches both inputs, and develops and writes print records onto an output magnetic tape.

7) KWICPRNT. This program receives the print record tape prepared via GENPRNTA, prepares the print records into a page of information, and prints the full report desired.

8) VLECK35U. This is a COBOL program designed to check the validity of the data in every 035 field and in every subfield "n" in a basic MARC record. In addition, an intermediate file of 035 fields and associated record numbers is built, sorted, and then sequenced to check for duplicate and/or omitted 035 data.

9) SKPGMU. This COBOL program is intended to create a 40-character sort key from specific tags or combination of tags in a MARC record. Each occurrence of a 245, 246, and 247 tag is located, and a sort key with the variable-field data found is built. If a 100, 110, or 111 tag is contained in the input record, the 245 sort key is split and the first 20 characters of the sort key will carry the lxx data, while the last 20 characters will carry the 245 data.

10) SKYCONV. This COBOL program is designed to perform two specific functions:

    a) to convert all lower-case characters in the sort key to upper-case for sorting purposes, and

    b) to sort the newly converted sort key in ascending order, using the COBOL sort verb.

11) PRTSKY. This COBOL program was developed to form a printed listing of the output of the SKYCONV program. This suite of programs (SKPGMU, SKYCONV, PRTSKY) was developed because BIBSKED would not provide quite the final form of output desired. The occurrence of 246 and 247 data in the output is indicated by a triple asterisk (***) code on the right side of the output page.

12) ARLMED. This COBOL program is intended to locate 52 machine-readable-detectable error conditions within each record in the master file. The conditions of test to be performed are built into the program and are not variable. The conditions tested include:

    a) fixed-field coding tolerances,

    b) incompatibilities between variable field presence and fixed-field coding,

c) variable-field presence related to variable-field presence, and

d) specific indicator and subfield coding irregularities in certain variable fields.

13) ARLMEDEX. This COBOL program is designed to identify any or all of 213 machine-detectable errors within the indicator and subfield coding of the variable fields of each MARC record. A comprehensive set of control cards is required; since control cards are used, values tested may be modified easily at the discretion of the user.

14) PRINTMED. This COBOL program accepts the two data sets created as a result of ARLMED and ARLMEDEX processing. The program sequences these data sets and then, with the aid of parameter cards, the error indicators are translated into readable text to be presented in an error listing.

Master File Data

Some idea of the characteristics of the data to be found in the master file can be gleaned from a study of the tables presented elsewhere in this report. These are by no means definitive in the sense that no other tabulations were possible; program capabilities were such that the tabulations that could have been produced were limited almost solely by imagination. Those given are intended to provide a kind of profile, albeit an incomplete and simplistic one.

CONCLUSIONS

Feasibility

A national serials data program appears to be both technically and economically feasible, economically feasible in the sense that cost-per-record at a production level was found to be reasonable. Such a program, however, must be conceived and executed as a separate entity, not as a part of something else. It is too important and too ambitious a project to be a mere appendage, and its objectives would not comport well with those of any single library.

The purposes, characteristics, and uses of such a program should be determined and clearly enunciated. While the program should be flexible enough to accommodate new conditions and needs, it should be firmly grounded so that its integrity is assured. It must be recognized that others are forging ahead in serials automation and that a national program, to be successful, must move quickly but carefully so that a superabundance of noncompatible mechanized data bases can be avoided.

A national data base can only be kept current with assistance from libraries in supplying the necessary information. To obtain such cooperation it would be desirable to provide libraries with a useful service or publications at the earliest possible date. In the meantime, efforts should be made to publicize the program. If a national serials data program is to be convincing and of maximum usefulness to the library community, it must be vigorous, well planned, well administered, and almost immediate. Unless a convincing effort is made, no truly national data base will result.

Authority File

The multiplicity of authorities used in building the NSPP file, coupled with the absence of a pilot project authority file, produced unsatisfactory results. The phrase "consistent data base" was often used to describe what was expected or at least desired. Yet no such data base could be constructed from numerous authority sources each of which was not only inconsistent with the others but was internally inconsistent as well. The need for an authority file became increasingly evident as the project progressed.

Rules of Entry

Terminology was often a stumbling block, as cataloging concepts appropriate to a card file were inevitably carried over to a machine-readable file.

Cataloging rules, after all, have been established for the creation of a specific tool, a card catalog, which has certain purposes and limitations. The computer, whatever it may be, is not a card catalog, and it is possible that new rules should be formulated and older rules modified to take fuller advantage of the computer's capabilities. So far, no one has seriously investigated, or even questioned, the suitability of existing cataloging rules in an automated context in which some card catalog limitations disappear but other problems arise.

This is true even of the MARC format, which was initially designed primarily for the computerized production of catalog cards. The phrase, "main entry," has unfortunate connotations in an automated environment. In a card catalog, a given work may be represented under several entries, one of which is the "main entry." But in a computerized file, there is only one entry and, to say the least, the phrase, "main entry," is redundant. That one entry may or may not correspond to the one a cataloger would choose as "main." The same line of reasoning can be applied to "cross references" and to "tracings." Since there is only one record, and it is of necessity untraced in the cataloging sense, these elements are unnecessary in a machine-readable file.

There are two key questions in an automated bibliography. 1) Can the item desired be accessed via the user's reference? 2) Can it be printed or displayed in the form desired? So long as both of these questions can be answered in the affirmative, it is of no more than passing concern to users of the output how the data are entered and stored in the computerized files. Naturally, it is a matter of considerable concern to systems designers that unnecessary repetition of bibliographic records be avoided and, since cost is a factor, they can generally be trusted to avoid the creation of extra records. If a proposed system produces the desired results at reasonable cost, the precise methods employed need not be a matter of concern.

Cost of Combined Upper and Lower Cases

Although no specific studies of the subject were made, it is estimated that the requirement for both upper and lower case roughly doubled costs at each stage of the project. Bibliographers noted for MARC editors which characters were to be capitalized; this required an extra character-by-character scan of verified data. MARC editors, similarly, were required to transmit this information without error to key-input operators. ATS operators were required to be especially alert for upper-case symbols. The comparison of input copy with keyed copy called for extra attention from the revisers. In addition, printing time was doubled. Also, the running of jobs was often delayed because there was only one printer equipped to output in both upper and lower case, and that one printer was available only at certain times of the day and for limited periods of time.

An added cost factor was the special programming required to output a printed list in alphabetical sequence, since the character codes for upper- and lower-case characters are necessarily different.

More important than any of these considerations, however, is the fact that the requirement for both upper and lower case actually resulted in more errors in the file. Some of these, to be sure, were the project's fault, but this requirement established the conditions for error to occur. Other errors came from the sources used, which were by no means consistent in such matters as capitalization.

From the viewpoint of both systems and cost, it would have been much better for the project to input and output in upper case only.

Expectations

Much had to be done before any results could be shown. Systems had to be designed, developed, and tested, possibly to be discarded in favor of others. Staff had to be interviewed, hired, and trained to do their jobs. The manifold mechanics of the operation had to be worked out in practice and changed when practice showed that theory was wrong. These were expensive and time consuming developmental tasks.

Inadequate allowance had been made for systems development, for the special difficulties in handling serials, and for major differences between a file created as a generalized data base and one created as an aid to localized control systems.

Decision making was the chief concern during the project's first three months. A month later, programming had been done, and four months after that the programs had been tested and de-bugged and could be considered operational.

Once these matters had been disposed of, it was possible to proceed with the assigned tasks. Even so, a pilot project, by its very nature, is a learning process and is usually intended to be such. It is to be hoped that the results of the project may ultimately serve as guides in the development of a national serials data program.

RECOMMENDATIONS

National Serials Data Program

It is recommended that the National Serials Data Program be continued and that a national serials data bank in machine-readable form be created. It is further recommended that the national serials data bank be planned to include not less than 200,000 titles by 1976.

Time is an important factor, because such a program must eventually involve the larger research libraries to some degree. Many of these libraries have already developed, are developing, or are planning to develop their own automated serials systems. Separately conceived and developed systems are unlikely to be compatible, and the existence of many incompatible systems, each very costly, would be detrimental to the development of a national system.

System

It is recommended that the National Serials Data Program employ a single computer and be in an on-line mode with it. During the first two to three years, this could be a shared facility, but it is likely that with the growth of the file and demands upon computer time the program will require its own computer. As for the on-line mode, this would permit direct access to the master file at all times, providing increased efficiency in updating, searching, and related activities. In addition, greater efficiency and reduced costs in programming and program maintenance could be expected to result from the use of on-line mode.

MARC Serials Format

It is recommended that the program employ the MARC serials format, since it is in the interests of the library and user communities to develop and use a single, standard format for the representation of bibliographic data.

Program Improvements

It is recommended that the programs be reviewed with the object of reducing both core requirements and running time. Core requirements could be reduced by partitioning the programs into smaller modules and by recoding the programs in assembly language. Running time could be reduced significantly by eliminating over-strikes (now used to produce a bold-face effect) in the printing and excessive double-spacing. Running time can also be reduced by providing an additional data set, together with appropriate programs, so that the frequency of full-file printouts may be significantly reduced. Cumulative listings of additions to the master file using this intermediate data set could be substituted. Finally, running time could be reduced substantially if output were largely (if not entirely) in upper case.

Authority File

It is recommended that the National Serials Data Program build its own machine-readable authority file. This authority file should contain the names of corporate bodies established according to the Anglo-American Cataloging Rules together with all associated or alternative forms of names. It should be interfaced with the master file on several levels. Names contained in the authority file should carry with them references to records in the master file associated with those corporate bodies. In addition, nonmain entry corporate names should be flagged with special codes in order to produce listings utilizing associated or alternative entries for particular requirements.

There should be no automatic updating of the authority file except in response to a specific input command. Similarly, although provision should be made for updating the master file via the authority file, there should be no automatic updating of the master file via the authority file except in response to a specific input command.

It is further recommended that the machine-readable authority file employ a basic MARC format, since the necessary interfacing with the master file could not otherwise be satisfactorily achieved.

It is also recommended that lists of additions to and changes in the authority file be published at regular intervals and distributed to all major libraries as a mutual aid. The libraries thus would be encouraged to report errors and changes to the national data bank while at the same time they would find the lists useful.

Input and Output

It is recommended that the need for upper- and lower-case output be thoroughly studied. If output can be in upper case only, many savings can be realized. Upper- and lower-case output is undeniably preferable, but it is not clear that it is necessary.

Rules of Entry

It is recommended that a small committee be appointed for the purpose of evaluating the applicability of cataloging rules to a machine-readable file. Each member of this committee should be thoroughly experienced in both cataloging and computer technology. The committee's objective should be to produce recommendations for the modification of existing cataloging rules. or the addition of cataloging rules, to exploit the capabilities of the computer more fully. This objective should be achieved with a minimum of change in existing cataloging rules.

MARC Serials Format Improvements

It is recommended that the innovations in the MARC serials format developed by the NSPP and described elsewhere in this report be made permanent features of the format.

It is further recommended that the Library of Congress give due consideration to the proposals for improvement in the format, presented in Appendix B.

BIBLIOGRAPHY

1). Information Dynamics Corporation. A Serials Data Program for Science and Technology. Prepared by William A. Creager and David E. Sparks. Reading, Massachusetts: 1965.

2). U.S. Library of Congress. Information Systems Office. National Serials Data Program: Phase I Final Report. Washington, D.C.: 1969.

3). Johnson, Donald W., and Richard L. Andrews. National Serials Pilot Project: Systems Documentation. Association of Research Libraries. Washington, D.C.: 1971.

4). Johnson, Donald W. National Serials Pilot Project: Manual of Procedures. Association of Research Libraries. Washington, D.C.: 1971.

APPENDIX A

SUGGESTIONS FOR FUTURE ORGANIZATION AND ADMINISTRATION OF A NATIONAL SERIALS DATA PROGRAM

1. The National Serials Data Program should be established as a separate agency within the Executive Branch of the federal government. The Department of Health, Education, and Welfare would seem to be appropriate.

Since it is to be expected that the National Serials Data Program would become relatively large in a fairly short span of time, it appears that it would be administratively much more sound to establish it as a separate agency rather than attach it to any existing agency. There would thus be far less competition for funds, staff, space, and equipment; moreover, in the event of a budget cut for a parent agency, the temptation to apply a disproportionate share of the cut to the National Serials Data Program would have been removed. In addition, establishment of a separate agency would constitute a much more permanent commitment on the part of the federal government. Finally, since this is an administrative proposition, it is only reasonable that such an agency be placed within the Executive Branch rather than the Legislative.

2. The National Serials Data Program should be governed by a board having five members: one designated by the Librarian of Congress, one designated by the director of the National Agricultural Library, one designated by the director of the National Library of Medicine, one designated by the board of directors of the Association of Research Libraries, and one designated by the president of the American Library Association. The chairman of the board would be elected by a majority vote of the board itself. (Terms of board membership and whether or not consecutive terms may be served would have to be determined.)

The board should be limited to a maximum of five members for the simple reason that boards and committees tend to become more and more inefficient as they are expanded beyond that size. Since the three national libraries will each have a very special interest in the National Serials Data Program, they should each be represented; the member libraries of the Association of Research

Libraries, on behalf of both themselves and of the library community, should be represented; and the American Library Association, as the most broadly based library association in the United States, should be represented.

3. The director of the National Serials Data Program would be directly responsible to the governing board through its chairman or other designated member.

4. The director of the National Serials Data Program would have sole administrative responsibility for the direction of the National Serials Data Program within the policy framework established by the board.

5. The administrative units responsible to the director of the National Serials Data Program will include the National Serials Data Bank and eight regional serials data banks.

It may be presumed to be inevitable that regional data banks will be created with or without the administrative blessing of the National Serials Data Program. Such regional data banks would not be able to correlate their operations and services in any truly effective way unless they are made a part of the National Serials Data Program and planned for accordingly. They should very definitely be limited in number. It appears that eight such data banks would suffice for the entire United States.

6. The director of the National Serials Data Program and his immediate staff would be located in the Washington, D. C., area.

7. The National Serials Data Bank should be situated somewhat to the west of the present population center of the continental United States.

Since each of the regional data banks would be in more or less continuous communication via computer-to-computer cable with the National Serials Data Bank, and since costs are a very definite factor, and since the volume of activity would relate more directly to population than to gepgraphic distance, and since the population center is moving westward, it is believed that the National Serials Data Bank would be best located a little to the west of the present population center of the United States. It should be located in a city large enough to be able to provide the specialized types of labor force and maintenance services. Denver comes to mind as a fairly obvious candidate, but other choices might be even better.

8. The regional serials data banks should be planned in detail at the same time as the detailed planning for the National Serials Data Bank is done, but they should not be created simultaneously. Rather, the regional serials data banks should be phased in one or two at a time.

If the detailed planning for the National Serials Data Bank and the eight regional serials data banks is done at the outset, adequate provision could be made for complete compatibility and, presumably, certain economies could be achieved by contracting for the necessary hardware and peripheral equipment at a single point in time but with various delivery dates.

9. Both the National Serials Data Bank and the eight regional serials data banks must be planned so that they will have both dedicated installations and dedicated circuitry.

Conceivably, the National Serials Data Bank could share a computer during the first year or two of its existence, with the provision that it would require a high and guaranteed priority on the use of the computer, but once the file size and the staff size acquired any sizable dimensions, it would be necessary to have an installation dedicated to its sole use. The same considerations apply to the establishment of regional serials data banks, except that each of them would be able to share a computer for a much more limited period of time owing to the fact that their files could be built up much more rapidly. Telephone circuits between data banks should also be dedicated in order to assure availability when needed and to minimize background noises emanating from other lines if on a shared basis.

10. Since the systems employed in the nine data banks would have to be facets of a single comprehensive system, it is implicit that all hardware configurations within the National Serials Data Program be completely compatible.

11. The National Serials Data Bank would include full bibliographic data on all serials. In addition, holdings would be recorded for each title for each of the three national libraries; in cases of titles not held by any of the three national libraries, an effort should be made to locate a reasonably complete file of that title in some United States library and to record that library's holdings of that title in the National Serials Data Bank. Furthermore, the fact that a given title is held in specified regional data banks would be recorded.

12. The amount of detail in the holdings records in the regional serials data banks will be greater and more explicit than in the National Serials Data Bank; on the other hand, the number of titles contained in any regional serials data bank will be very substantially less than in the National Serials Data Bank.

It is obvious that the degree of detailed information to be found on national, regional, and local levels will differ. So far as recording of holdings is concerned, the most complete and current information will always be on the local level; so far as relative completenessof the file is concerned, this can be achieved only on the national level. It follows that the regional files would fall between these two extremes.

13. The National Serials Data Bank will be the authority for all regional serials data banks with respect to bibliographic data, format, general systems, etc. It is expected, however, that information for the updating of the National Serials Data Bank would in many instances flow from the regional serials data banks to the National Serials Data Bank and, after appropriate verification and input, revised data would flow back from the National Serials Data Bank to the regional ones.

14. The National Serials Data Bank would accept and process queries received from any of the three national libraries or any of the regional serials data banks, but only from these sources. The regional data banks would receive and process queries from either participating libraries in their respective regions or from the National Serials Data Bank.

15. The most detailed and most current records with respect to holdings would be those maintained in and by individual libraries.

16. The National Serials Data Bank should function also as a switching center, routing queries from a given regional serials data bank to another regional bank when circumstances warrant.

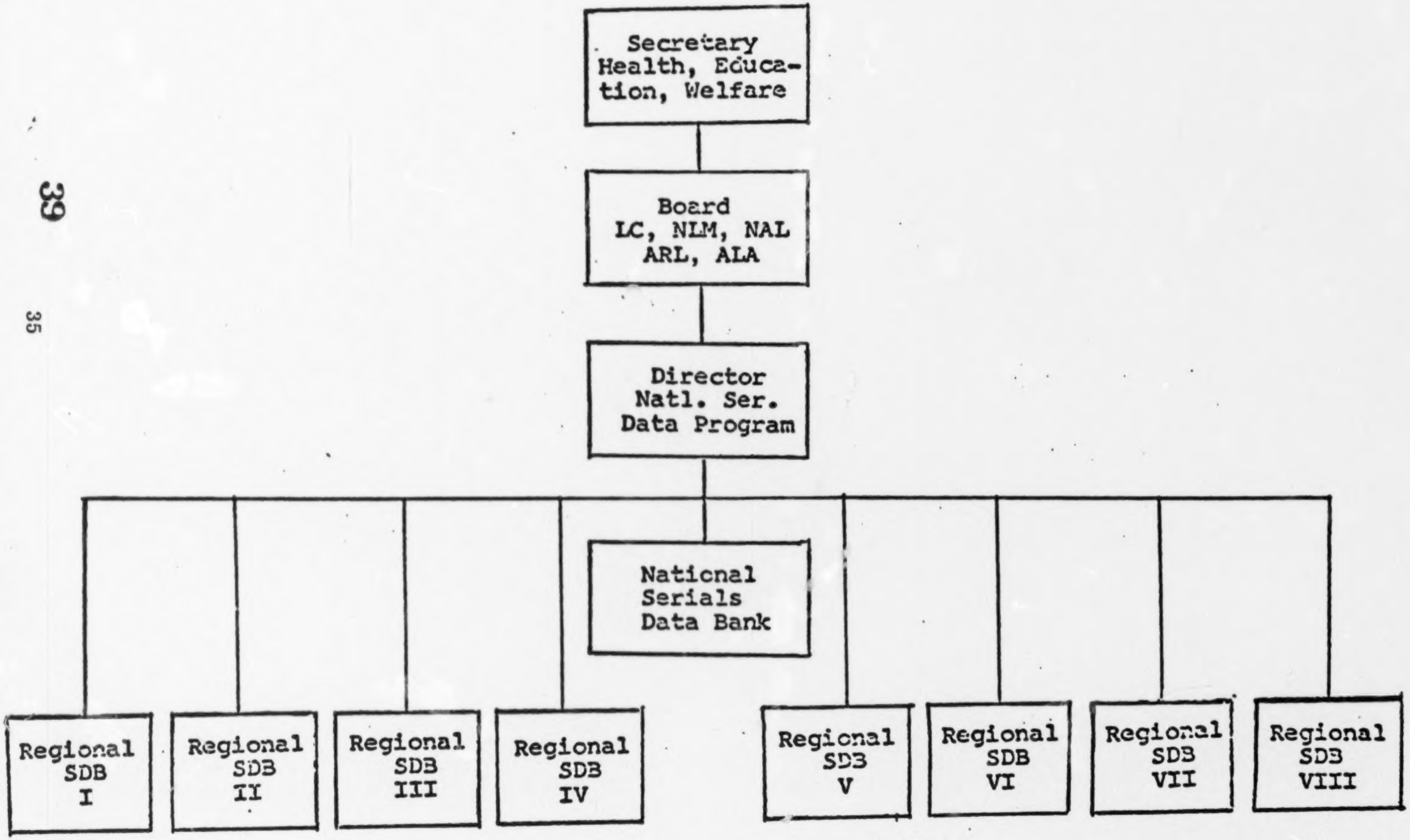
It is believed that direct connection between the regional serials data banks would unnecessarily add to the operational costs. Indirect connections via the National Serials Data Bank would suffice and should be less costly.

17. The National Serials Data Bank will record in its files both Standard Serial Numbers and International Standard Serial Numbers according to procedures to be determined by the Library of Congress.

18. It is essential that continuous updating procedures be established as part of the system and that updated information flow up, down, and across the network.

CHART II

NATIONAL SERIALS DATA PROGRAM
PROPOSED ADMINISTRATIVE
ORGANIZATION CHART


Secretary Health, Education, Welfare
Board LC, NLM, NAL ARL, ALA
Director Natl. Ser. Data Program
National Serials Data Bank
Regional SDB I
Regional SDB II
Regional SDB III
Regional SDB IV
Regional SDB V
Regional SDB VI
Regional SDB VII
Regional SDB VIII

APPENDIX B

SUGGESTED IMPROVEMENTS IN THE MARC SERIALS FORMAT

1) Fixed Fields

   a) Avoid negative coding wherein the absence of data acquires a positive meaning.

   b) Provide additional codes for boxes 11 and 12 to allow for maps.

   c) Assign codes to box 23 on some basis other than political status at time of initial input. It might even be a good idea to avoid country codes, as such, entirely, utilizing instead codes representative of physical location.

   d) Define for users of the format what dates are to be used in boxes 21 and 22 in the case of reprints of serials; at present this is unclear.

2) Variable Fields

   a) Field 041 (languages). Instead of the present method of showing the language from which translated, present such data in an additional subfield.

      Add "mul" to the list of valid language codes for this field.

      Provide for a subfield "a" over-ride with field 041 so that it will not be necessary to repeat language of text when the field is being used only to show languages of abstracts, summaries, or original publication.

   b) Field 100 (main entry, personal name) should have a subfield for relator, such as is provided in field 700. This should probably be subfield "e."

c) Field 245 (full title). Make subfield "a" repeatable to allow for those instances where the title of the work is presented in two or more languages.

Subfields "b" and "c" should also be made repeatable for the same reason as above, but repetition of either subfield "b" or subfield "c" should be invalid unless subfield "a" has been repeated.

Add a subfield to field 245 for edition, as it would be more logical, more convenient, and would reduce file size as opposed to the use of a separate field (field 250) for such data.

d) Field 246 (varying forms of title). Make adjustments in the meanings of indicators as required by the recommended changes in the use of field 245.

e) Field 247 (former titles and their variants). Eliminate this field, as it only partially serves the purpose for which it was intended. As a substitute, create a new field (perhaps 781) for antecedent titles more than one generation removed from the title shown in field 245.

f) Fields 710 and 711 (others associated with the work--corporate bodies). Provision should be made for separate subfield encoding of subtitles, sections, etc., so that titles as shown in these fields can be correlated more readily with titles as shown in field 245.

g) Linking entries, general. Some additional fields are needed, e.g., to link between a serial that contains only abstracts of a translation and both the translation and the original work.

Provide for separate subfield encoding of subordinate units of corporate bodies so that machine correlation of corporate names appearing in linking-entry fields with the same bodies in fields 110 and 111 can be more readily accomplished.

Provide for linking to some volumes only of a serial.

Provide an additional field for "war merger" types of relationships. The present linking entries are not really adequate to express this sort of thing, especially since in many cases the so-called "war merger" resulted in no actual merger and was no more than a device to preserve an artificial continuity.

h) Field 765 (original language entry). Make this field repeatable, since there are some serials that are translations of more than one serial.

i) Fields 770, 772, and 777 (supplement entry, parent record entry, and "issued with" entry). Provide a means of showing dates of relationship, and perhaps volumes as well in these fields. At present only the beginning date of a relationship may be shown.

j) Field 780 (preceding entry). The secondary indicators 4 and 7 are incompatible with the nonrepeatability of the subfields in this field, since the data for each of these indicators require that two titles be input. Either the indicators 4 and 7 should be eliminated, or the subfields should be made repeatable.

k) Field 785 (succeeding entry). The secondary indicators 6 and 7 are incompatible with the nonrepeatability of the subfields in this field, since the data for each of these indicators require that two titles be input. Either the indicators 6 and 7 should be eliminated, or the subfields should be made repeatable.

l) Field 850 (library code and holdings) should be made repeatable. In addition, program changes should be made that will permit the computer to treat subfield "a" as an extension of the tag so that proper interfiling of holdings of various libraries can be accomplished automatically rather than by an improper application of site numbers.

APPENDIX C

SUMMARY OF NATIONAL SERIALS DATA PROGRAM

Introduction

This outline features the more important provisions of the Technical Development Plan of the National Serials Data Program and revises and brings up to date some subsequent agreements particularly in the area of administrative arrangements. The plan, which will serve as the basis for the further development of the national serials cooperative effort, is concerned primarily with the detailed technical steps, many necessarily experimental.

This document will provide the program staff with the policy decisions essential for the current phase. Within the policies set by this document, the director and staff of the National Serials Program will decide upon and implement techniques and procedures. The director of the program will have full authority for day-to-day operational decisions and will have the responsibility for reporting periodically to the directors of the three national libraries through the Librarian of Congress.

Need for the Program

Although the importance of concentrating efforts and resources on finding more economical and satisfactory ways of controlling access to serial literature in the research library community is obvious, the need for a Serials Data Program goes even deeper than the need for control of serial litarature for reference and research purposes. The problems this program must face and attempt to solve are problems fundamental to the economics of American librarianship. Librarians have long been concerned with the need for a central source of serial cataloging information and the need for an economically feasible system of handling serials that would eliminate a considerable number of the costly duplicative input and conversion projects existing today. This is not to be construed to mean, however, that this project can be expected to furnish complete and detailed cataloging information.

The Council on Library Resources recognized these needs in its Annual Report for 1970, which states that "What is expensive is the cataloging necessary to identify serials. A National Serials Program should strive to eliminate some of the duplication of cataloging effort now going into this class of materials."

Most of the steps that have been taken in the last one hundred years toward standard bibliographic description and centralized cataloging have been book or monograph oriented. The Higher Education Act of 1965 and the subsequent Shared Cataloging Program at the Library of Congress are signal achievements in the area of standards and centralized cataloging, yet this program is essentially geared to and limited to monographic publications.

Present machine-readable files of serials information range from sketchy records containing minimal information to the augmented catalog of Project INTREX. A survey made by the Library of Congress in 1968 showed that approximately 350 institutions had some of their serials information in machine-readable form. These data bases have been developed independently and have involved substantial duplication of costs and efforts. Continuation of this unnecessary duplication of effort should not be necessary for American libraries.

The adoption in late 1970 of the Standard Serial Number (Z 39.9) presupposes a general agreement on the need for unique identification of titles in any serial system. It is essential that the National Serials Data Program assume the responsibility for the assignment of SSNs and for the dissemination of information about SSNs.

The need for a standardized, machine-readable data bank is urgent. A vast amount of duplication of effort and resources can be avoided once records with standardized bibliographic description are converted and the data communicated to other libraries and institutions via a standard communications format.

Objectives

The major objectives of the National Serials Data Program are:

1) to provide U.S. libraries, including the three national libraries, with an authoritative bibliographic resource upon which serials processing systems can be built;

2) to provide a base record of serial titles to which the Standard Serial Number can be permanently affixed, thus ending most of the confusion about precise identification of serials; this file of records will serve as the U.S. Register of Standard Serial Numbers;

3) to provide a bibliographic resource for serials which will supply important cataloging information to libraries and at the same time will permit the uniform transfer of data on serials among libraries;

4) to provide a base from which several kinds of library tools can be printed; and

5) to provide a serial system which will constitute the U.S. segment of the developing International Serials Data System.

Policy

The principles of basic policy which follow are based on the findings and recommendations of the National Serials Pilot Project, as well as on agreements previously reached within the Task Force. These principles have the concurrence of the Executive Director, Association of Research Libraries and of the President, Council on Library Resources--two organizations very intimately involved in the pilot project. These policy matters concern:

1) Authority File. A machine-readable authority file is requisite to the success of any large serials file which is to be automated. This requirement is absolute when differing forms of corporate entry exist in the source files and it is not possible to recatalog them into a consistent file. This describes the existing national serials data base accurately. Accordingly, a machine-readable file will be built which will include, as a minimum, all corporate entries appearing in the file, showing the form of each as used by each of the three national libraries. The ultimate aim is to have a single authoritative entry under AA rules. The system should have a built-in capability to output whichever form of corporate entry is desired for whatever purpose; this should include the capability to use the AA form of entry against the time in the future when this becomes acceptable to all three libraries.

2) Data Base. It is proposed to limit the scope of initial coverage to finite segments of the literature. The National Serials Data Program will strive at first for coverage of new titles in all fields and of retrospective titles in the fields of science and technology. Initial sources of retrospective titles will be the pilot project data base, NLM titles, NAL titles, and LC Science and Technology Reading Room titles. Eventually, all live scientific and technical titles in the three national libraries will be included.

3) Data Elements. In view of the large size of even a selected data base, it is necessary to review carefully the estimated number of data elements which can be handled effectively by this system; sample studies have indicated that an average record of about twelve data elements will cover most needs.

Within the framework of a data record of this average size, it is proposed to develop a system that will supply enough cataloging data to eliminate a good deal of the duplicative cataloging effort now being performed in thousands of libraries, as well as a record which will provide a limited range of location information.

In Phase III a comprehensive base of records will be developed which will accurately and uniquely identify serial titles, which will provide certain important data elements for serials cataloging purposes, and which may also be used to locate journals. It is recognized that this record will not satisfy every library's requirement; librarians desiring additional elements for their individual internal purposes can build on these basic elements.

4) A separate effort is required. Everyone concerned with the pilot project is firmly convinced that the development of the National Serials Data Program can proceed satisfactorily only if it is kept separate from the internal serials operations of any library. This is so for several reasons:

   a) The total requirement for internal operations of any major library cannot be met by the abbreviated record here proposed; the program cannot afford a longer one.

   b) Since the National Serials Data Program will transcend the needs of any of the individual libraries, it is essential that the staff developing and implementing this program be separate from and independent of the operating staff of any of the individual libraries. A successful program requires that those responsible for it have authority to make decisions independently.

5) MARC Serials Format. The MARC serials format will continue to be used in the National Serials Data Program.

6) Expanded Character Set. The use of an expanded character set, particularly the use of upper and lower case, will be continued.

7) Administrative Arrangements. The director will operate the program in accordance with the Technical Development Plan. He will report to the directors regularly through the Librarian of Congress. He will also keep the Task Force informed of progress and problems through the chairman of the Task Force.

The Task Force is expected to advise and assist the director in the progress of the program and the prompt resolution of program problems including suggested amendments to the Technical Development Plan as needed. Program problems which cannot be resolved promptly will be brought to the attention of the three library directors by the program director for resolution. Decisions and directions on the program will be given to the program director by the Librarian of Congress.

Similar procedures will be followed on amendments to the Technical Development Plan and the review and approval of annual program and budget plans. The program will begin operation on April 17, 1972, at a level of $2,000 per month per library. The program will continue in fiscal 1973 at this level, perhaps supplemented by funds from other sources. A program plan and budget for fiscal 1974 will be prepared for review by August 14, 1972, so that each library can include the agreed-upon portion in its fiscal 1974 appropriation request.

Organization and Reporting Channels

Because of the limited funds available, the staff of the National Serials Data Program must be kept small. The minimum number with which a reasonable amount of progress can be made is something near the staff level of the pilot project, which had seven full-time people. It should be pointed out, however, that at least for the first year, the level of serials expertise required to solve operational and technical problems of the program will be high. For this reason, average salaries for this first year will be high. However, once the initial problems are solved, subsequent hiring will be carried out at lower entrance levels.

The program must have a director; it must have serials librarians; and it must have clerical support for input typing. Experience in the pilot project gives clear evidence of a requirement for at least one systems analyst and one programmer. Since these skills cannot be contributed by the national libraries, they must be provided for in the program budget.

The director of the program will report to the Librarian of Congress. Advice and assistance will be furnished by the Task Force through the chairman. The assistance and advice of the advisory group, originally appointed to guide the pilot project, will also continue.

Current Operations

In order to maintain the momentum gained during the pilot project, work must continue without interruption. Most of the professionals on the project staff are now employed by the Library of Congress; two are being used in the continuing program. Enough has now been learned to make quite clear the major direction of this project. It must not be permitted to lapse into another period of theoretical study. The National Serials Data Program will never have a better opportunity than now to "get off the ground." Its functions are currently limited to the primary purpose of identifying and recording a body of bibliographic information on serials. A secondary objective is the recording of simplified holdings information, beginning with those of the three national libraries, with gradual expansion to include data from other major research libraries. A major by-product of immense value will be the machine-readable authority file. All files, techniques, procedures, and programs developed by this system will be made freely available to the library community.

Library of Congress
March 1972

APPENDIX D

OPERATIONS FLOWCHARTS

List of Flowcharts

A, 1 Verification (LC Official File)
A, 2 Verification (LC Serial Record)
A, 3 Verification (New Serial Titles)
A, 4 Verification (ULS)

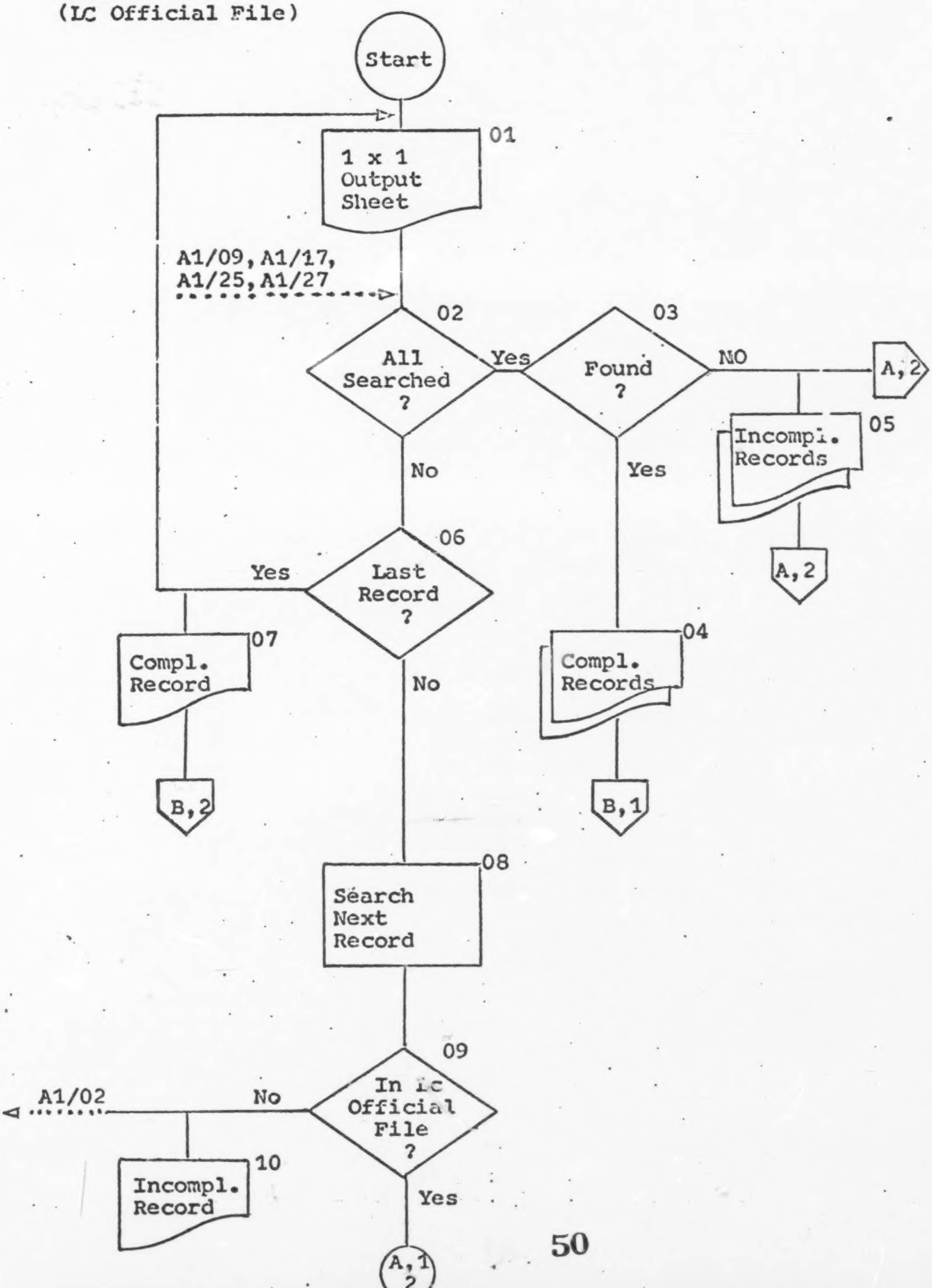
B, 1 Editing
B, 2 Editing--Supplementary Search

C, 1 Key Input
C, 2 Key Input--Revision
C, 3 Key Input--Correction

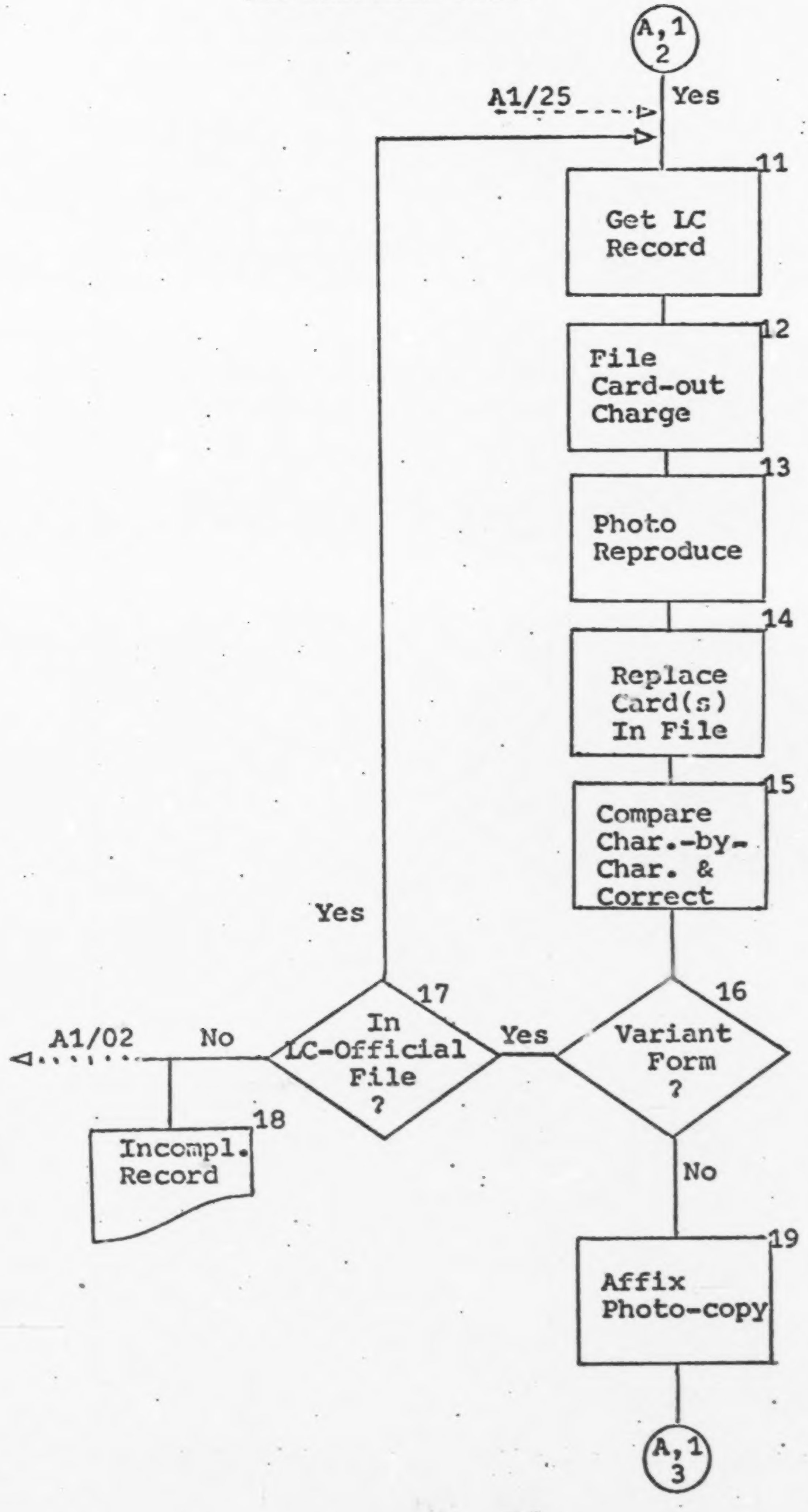
D, 1 Update-run--'ATS'
D, 2 Update-run--'Bibsys'
D, 3 Update-run--'New Master'

E, 1 Master-file Corrections (ATS Diagnostics)
E, 2 Master-file Corrections (Maintenance Diagnostics)
E, 3 Master-file Corrections (Vleck35U)
E, 4 Master-file Corrections (MED)

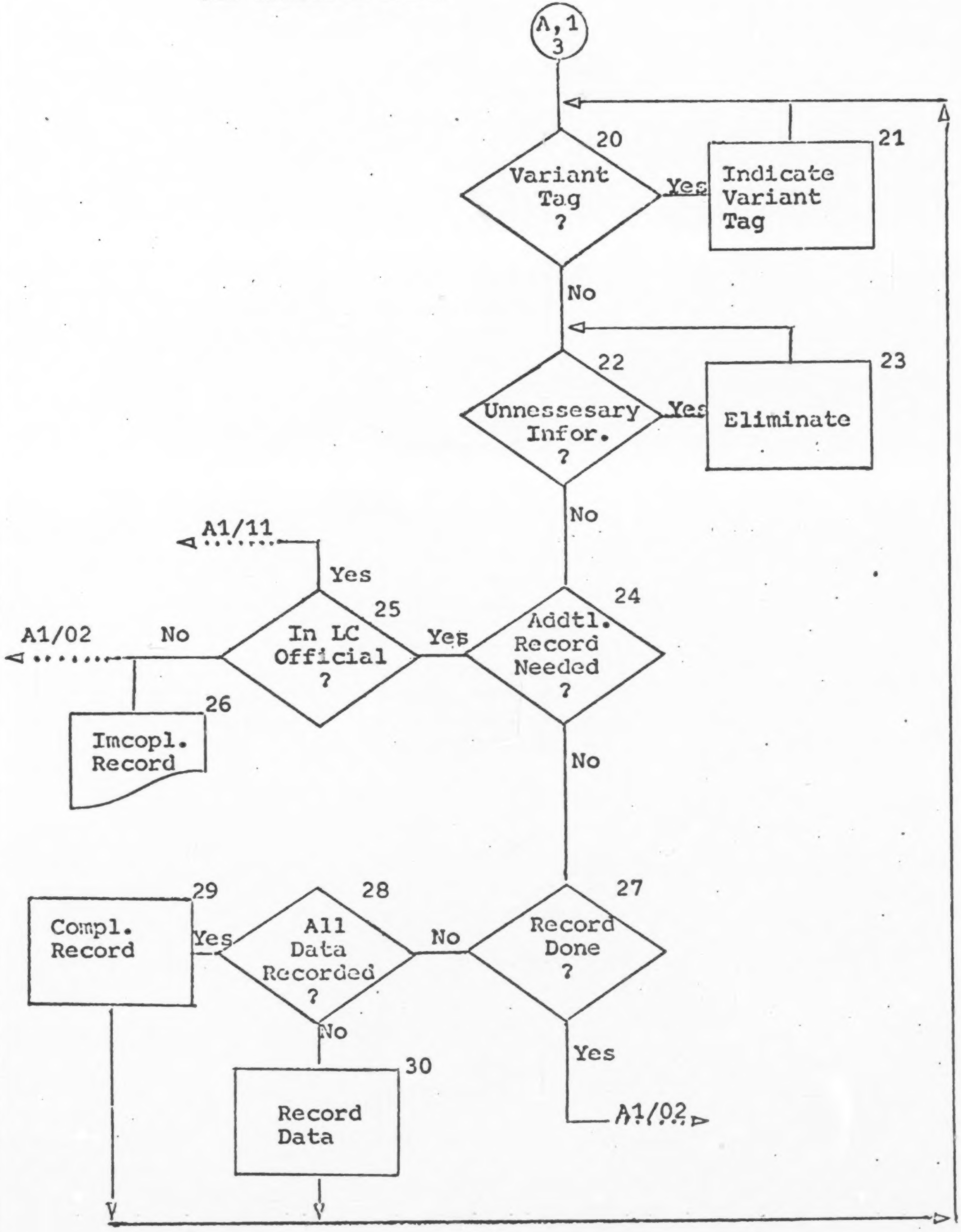
NSPP - Operations
A,1 - Verification
(LC Official File)


Start
01
1 x 1
Output
Sheet
A1/09, A1/17,
A1/25, A1/27
02
All
Searched
?
Yes
03
Found
?
NO
A,2
05
Incompl.
Records
A,2
Yes
No
04
Compl.
Records
B,1
06
Yes
Last
Record
?
No
07
Compl.
Record
B,2
08
Search
Next
Record
09
A1/02
No
In LC
Official
File
?
10
Incompl.
Record
Yes
A,1
2

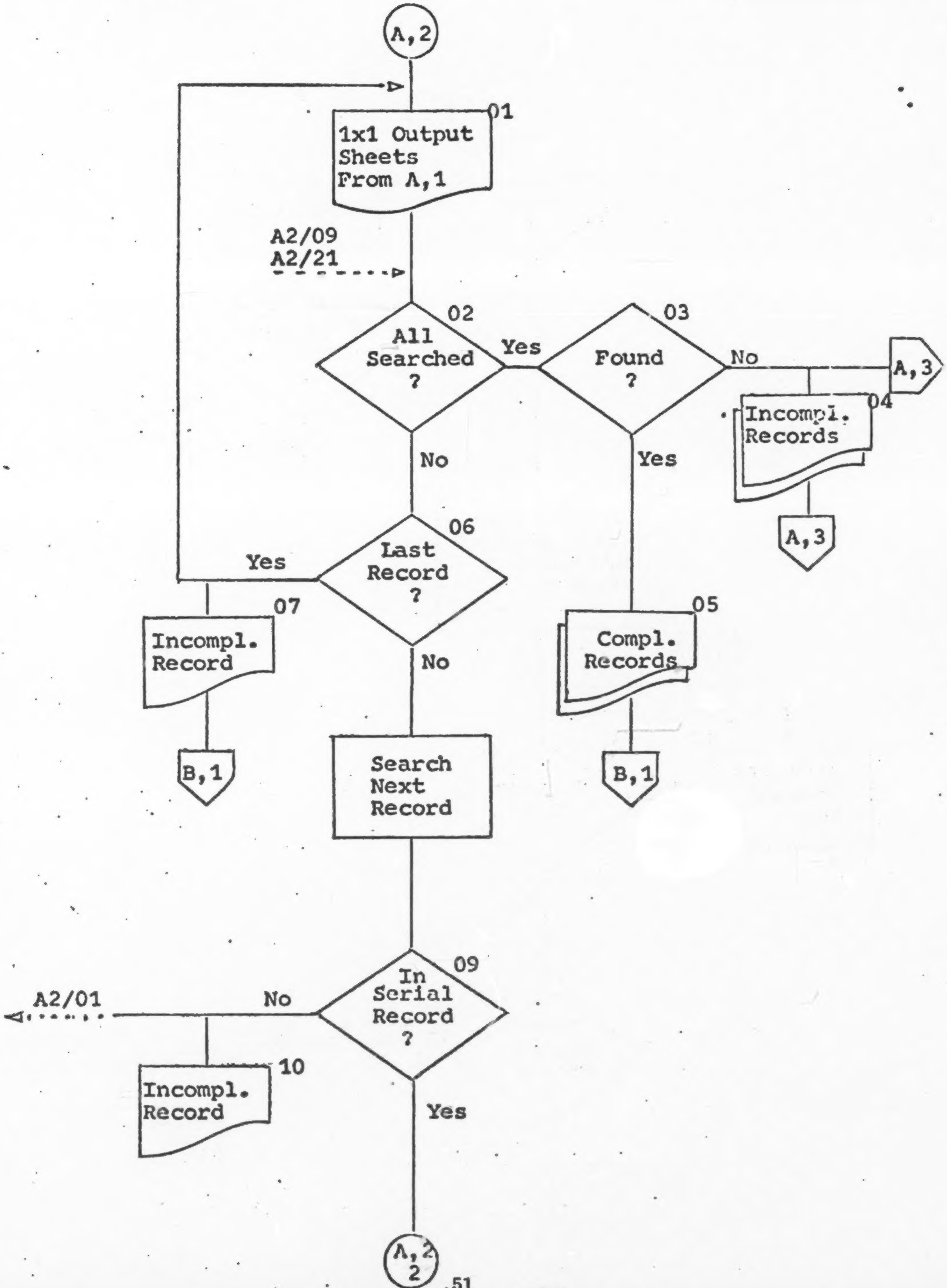
NSPP - Operations
A,1.2 - Verification
(LC Official File)


A,1
2
A1/25
Yes
11
Get LC
Record
12
File
Card-out
Charge
13
Photo
Reproduce
14
Replace
Card(s)
In File
15
Compare
Char.-by-
Char. &
Correct
Yes
17
In
LC-Official
File
?
16
Variant
Form
?
A1/02
No
Yes
18
Incompl.
Record
No
19
Affix
Photo-copy
A,1
3

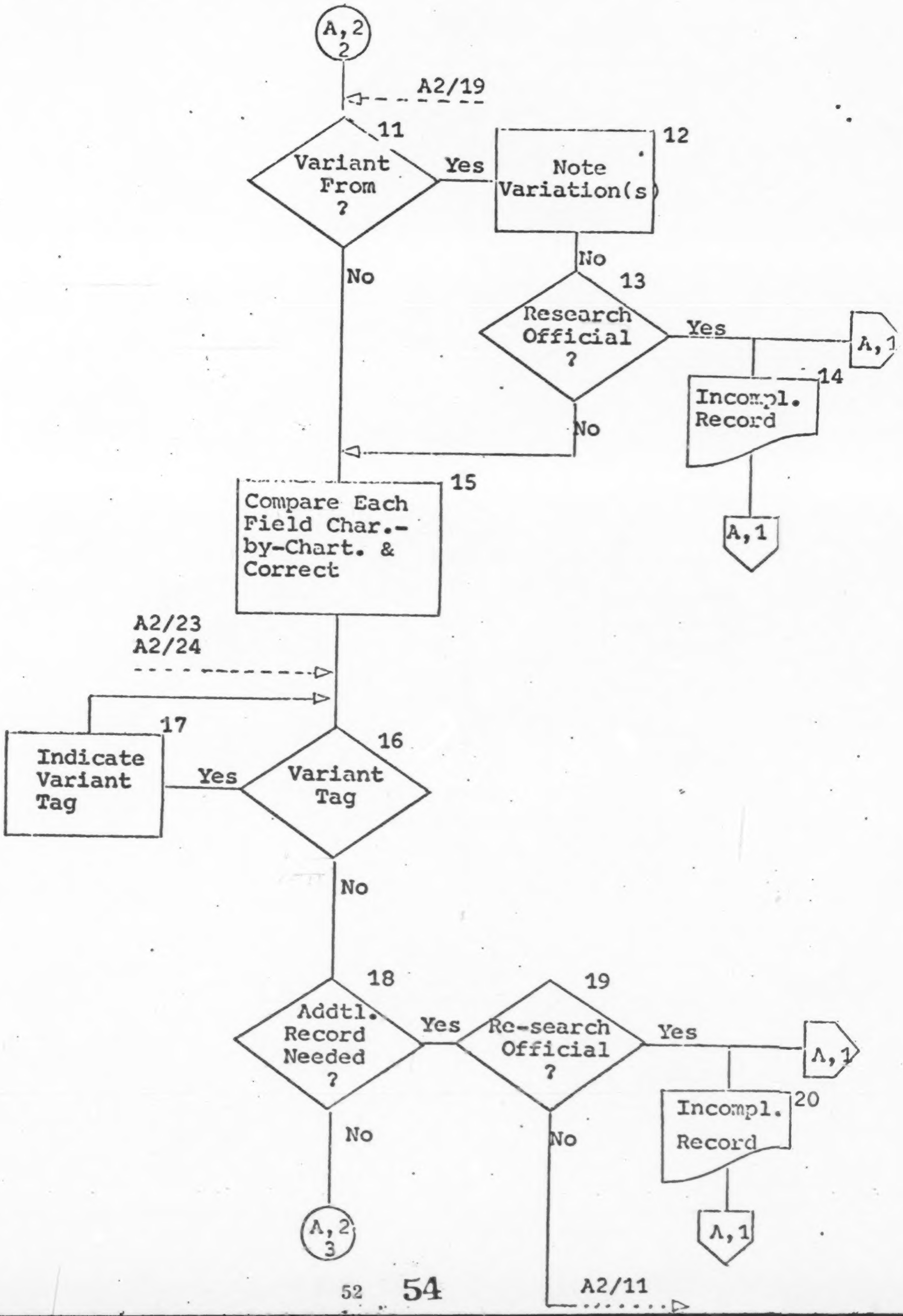
NSPP - Operations
A,1.3 - Verification
(LC Official File)


A,1
3
20
Variant Tag ?
Yes
21
Indicate Variant Tag
No
22
Unnessesary Infor. ?
Yes
23
Eliminate
No
24
Addtl. Record Needed ?
Yes
25
In LC Official ?
Yes
A1/11
No
A1/02
26
Imcopl. Record
No
27
Record Done ?
Yes
A1/02
No
28
All Data Recorded ?
Yes
29
Compl. Record
No
30
Record Data

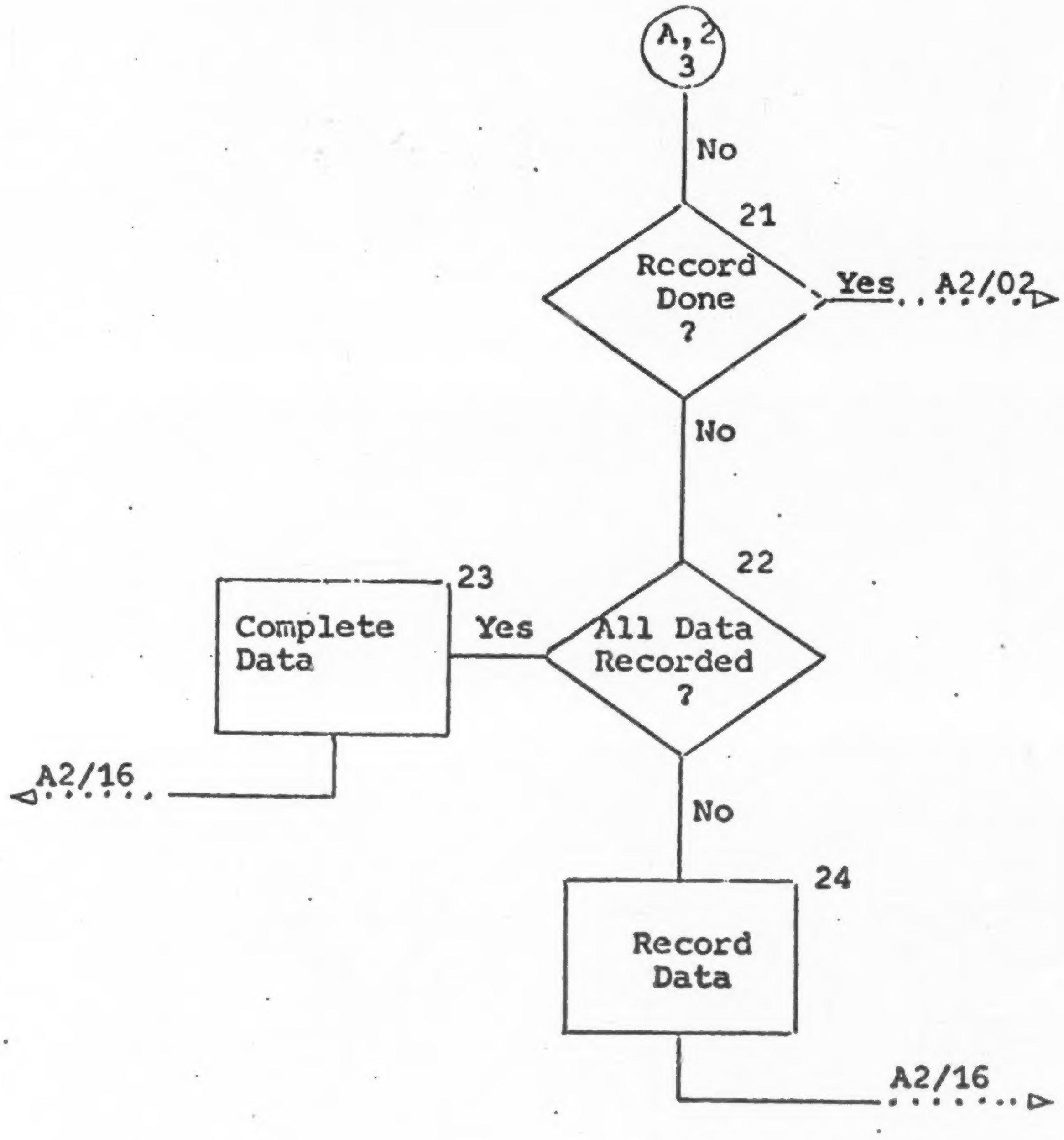
NSPP - Operations
A,2 - Verification
(LC Serial Record)


A,2
01
1x1 Output Sheets From A,1
A2/09
A2/21
02
All Searched ?
Yes
03
Found ?
No
A,3
04
Incompl. Records
A,3
Yes
No
06
Last Record ?
Yes
07
Incompl. Record
B,1
05
Compl. Records
B,1
No
Search Next Record
09
In Serial Record ?
No
A2/01
10
Incompl. Record
Yes
A,2
2

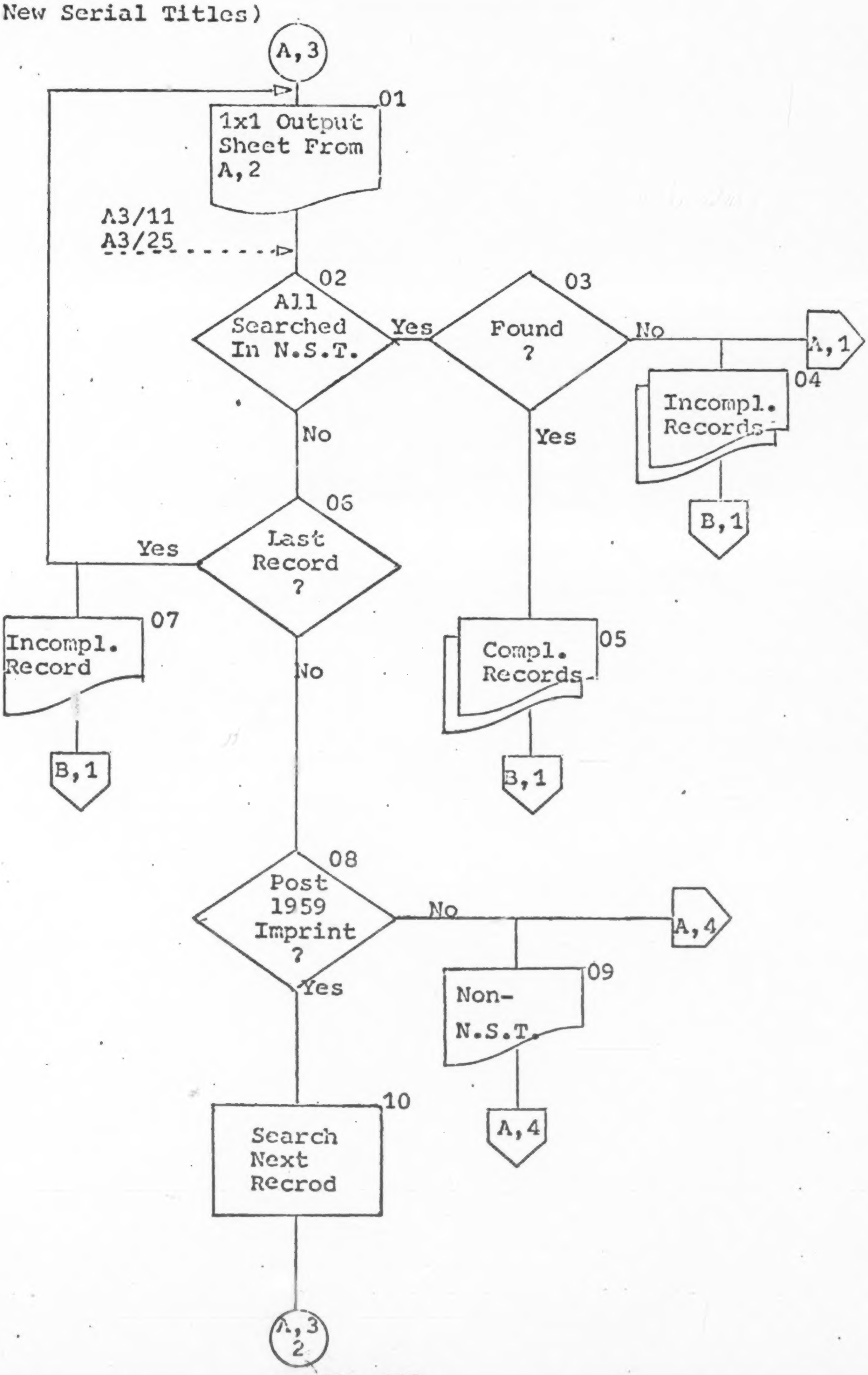
NSPP - Operations
A,2.2 - Verification
(LC Official File)


A,2
2
A2/19
11
Variant
From
?
Yes
12
Note
Variation(s)
No
13
Research
Official
?
Yes
A,1
14
Incompl.
Record
A,1
No
No
15
Compare Each
Field Char.-
by-Chart. &
Correct
A2/23
A2/24
17
Indicate
Variant
Tag
Yes
16
Variant
Tag
No
18
Addtl.
Record
Needed
?
Yes
19
Re-search
Official
?
Yes
A,1
20
Incompl.
Record
A,1
No
No
A,2
3
A2/11

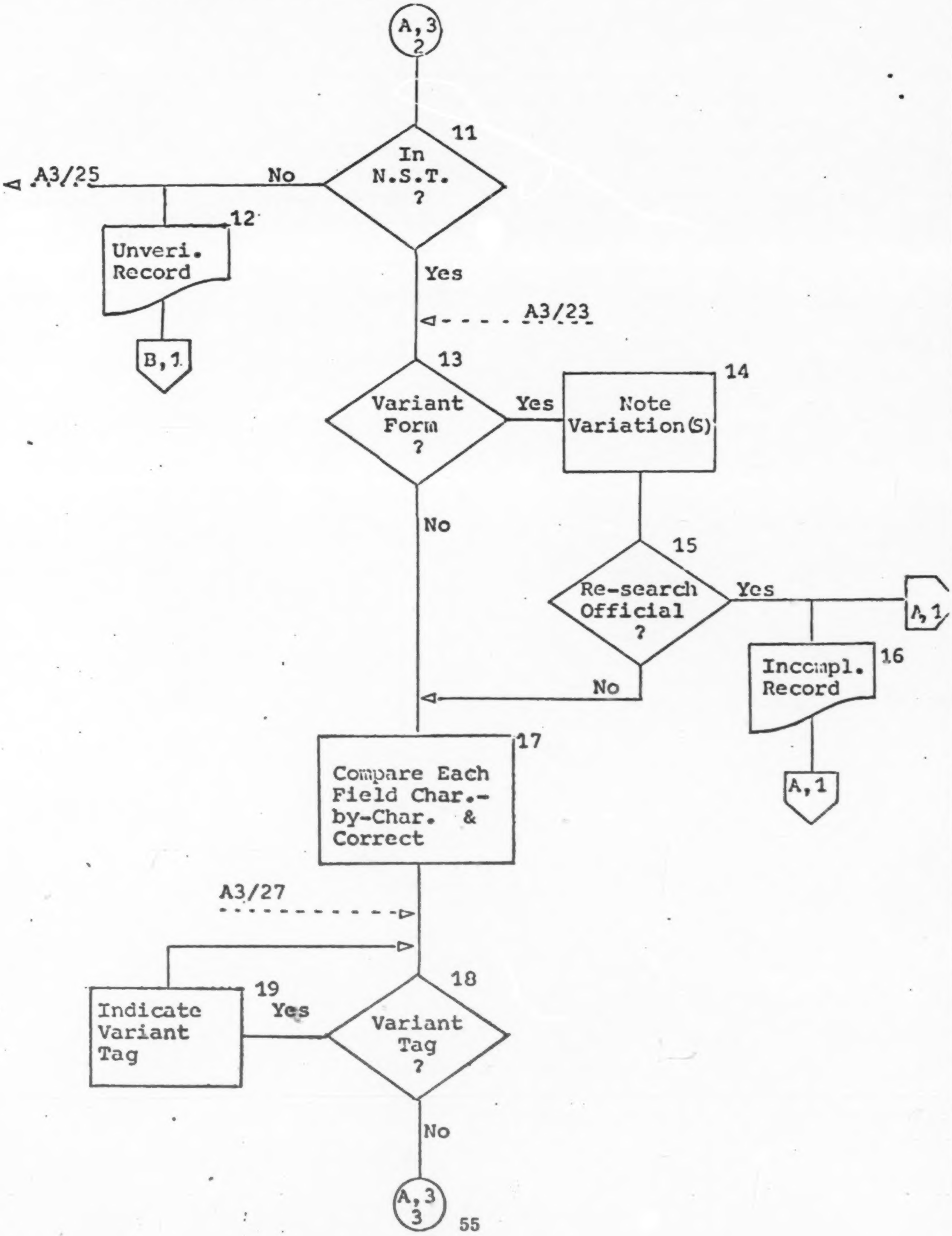
NSPP -- Operations
A,2.3 - Verification
(LC Serial Record)


A,2
3
No
21
Record Done ?
Yes
A2/02
No
22
All Data Recorded ?
Yes
23
Complete Data
A2/16
No
24
Record Data
A2/16

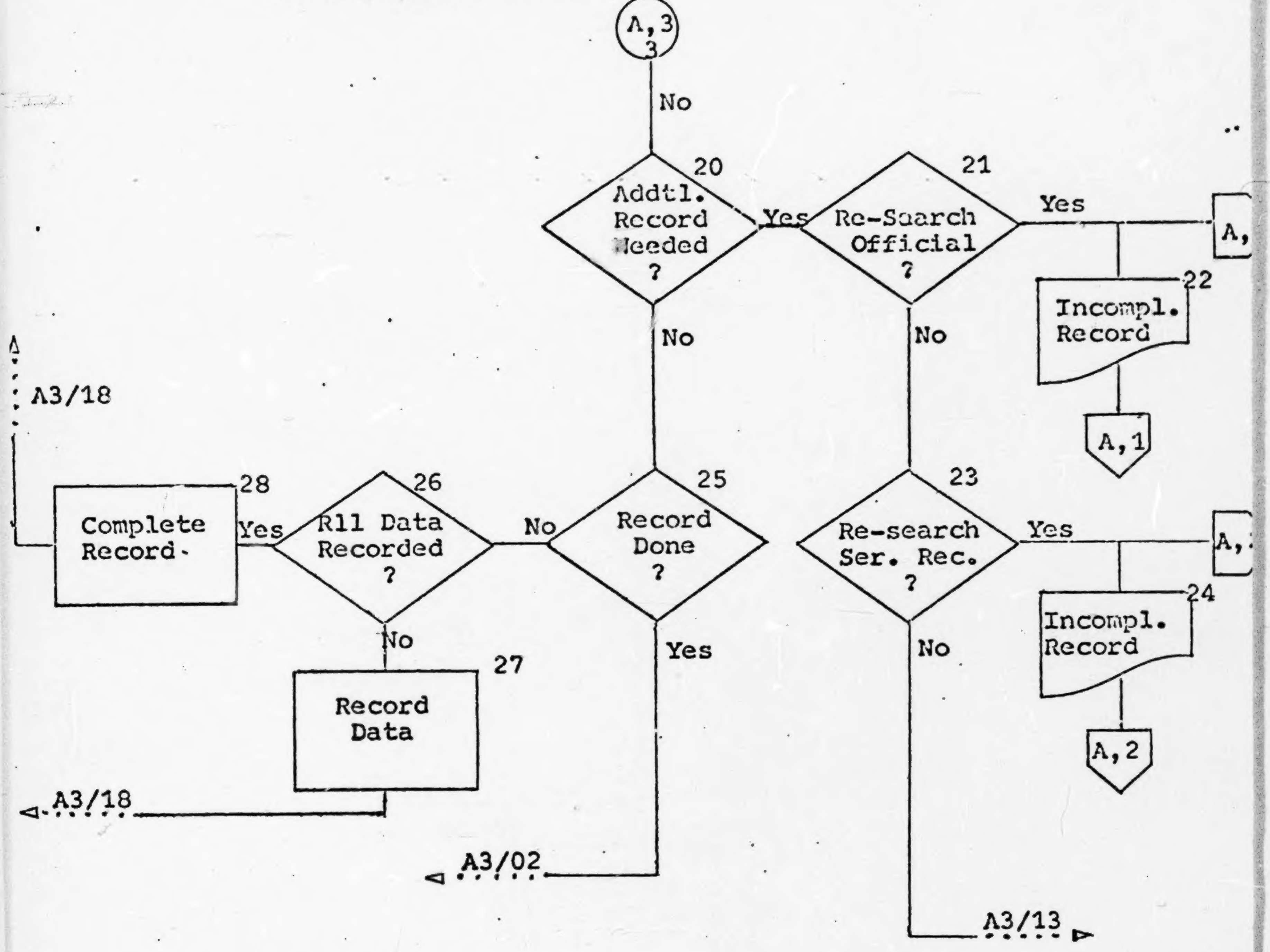
NSPP - Operations
A,3 - Verification
(LC New Serial Titles)


A,3
01
1x1 Output Sheet From A,2
A3/11
A3/25
02
All Searched In N.S.T.
Yes
03
Found ?
No
A,1
04
Incompl. Records
B,1
Yes
No
06
Last Record ?
Yes
07
Incompl. Record
B,1
No
05
Compl. Records
B,1
08
Post 1959 Imprint ?
No
A,4
09
Non-N.S.T.
A,4
Yes
10
Search Next Recrod
A,3 2

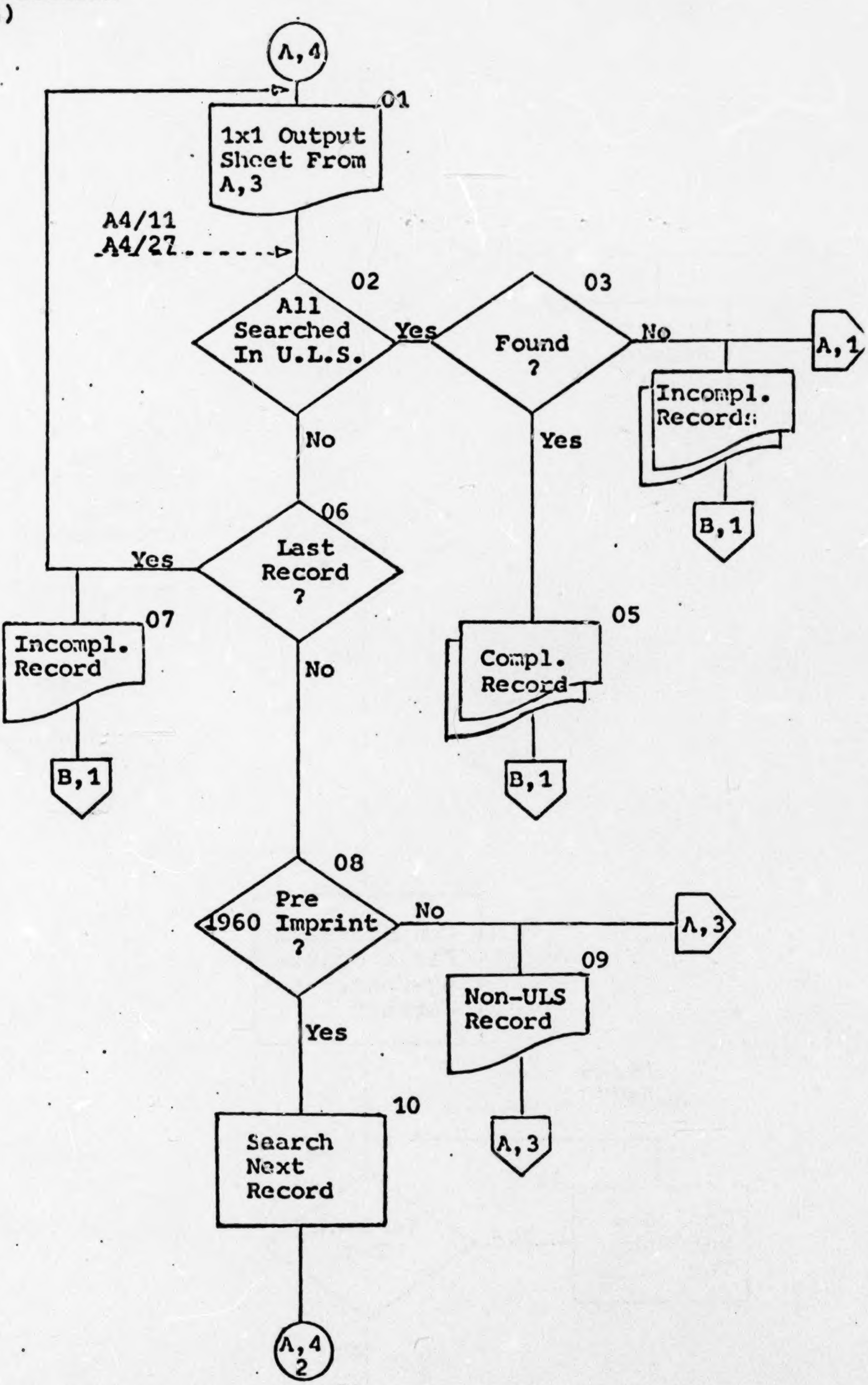
NSPP - Operations
A,3.2 - Verification
(LC New Serial Titles)


A,3 2
11
In N.S.T. ?
No
A3/25
12
Unveri. Record
B,1
Yes
A3/23
13
Variant Form ?
Yes
14
Note Variation(S)
No
15
Re-search Official ?
Yes
A,1
16
Incompl. Record
A,1
No
17
Compare Each Field Char.-by-Char. & Correct
A3/27
19
Indicate Variant Tag
Yes
18
Variant Tag ?
No
A,3 3

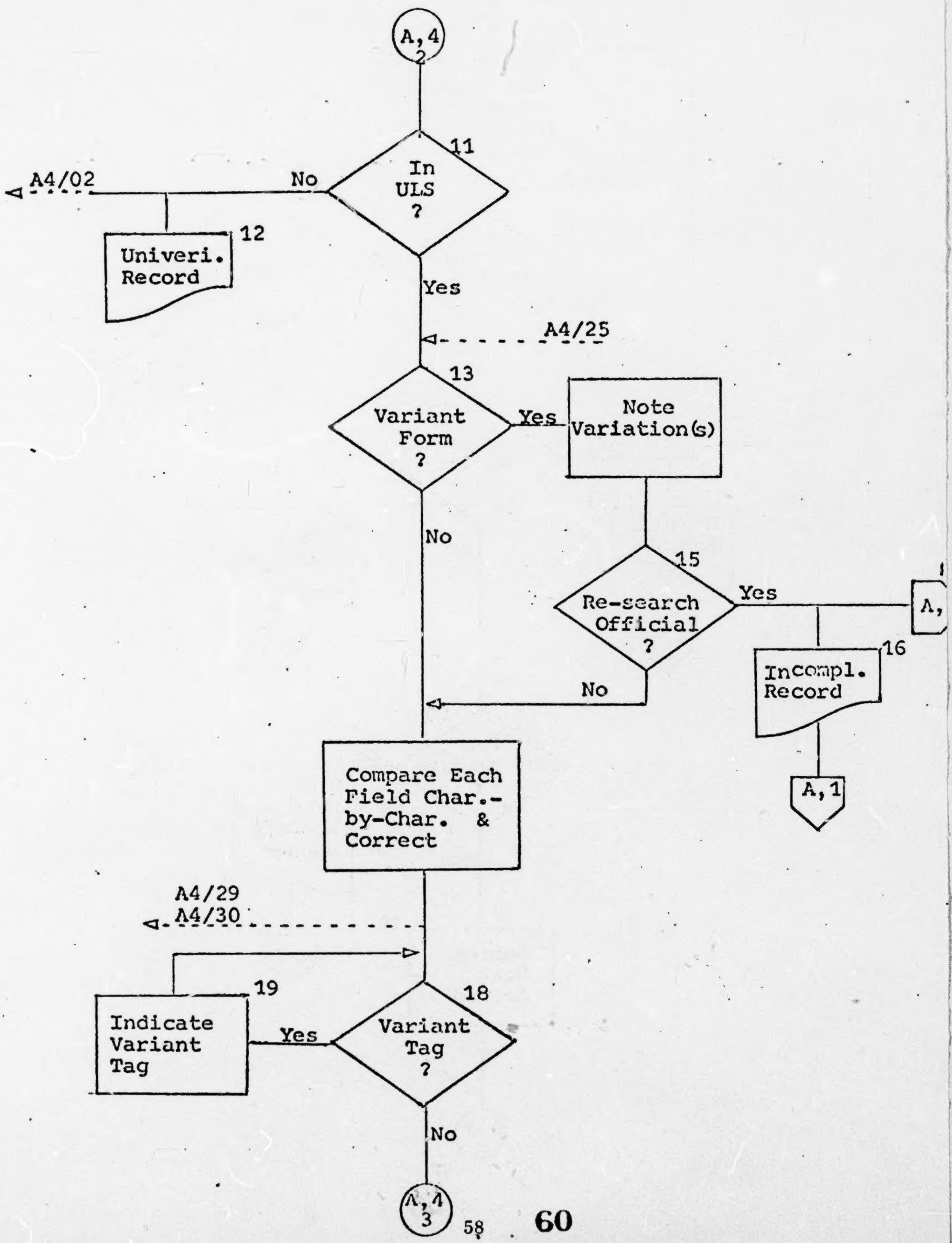
NSPP - Operations
A,3.3 - Verification
(LC New Serial Titles)


A,3
3
No
20
Addtl.
Record
Needed
?
Yes
21
Re-Search
Official
?
Yes
A,
22
Incompl.
Record
A,1
No
No
A3/18
28
Complete
Record
Yes
26
All Data
Recorded
?
No
25
Record
Done
?
23
Re-search
Ser. Rec.
?
Yes
A,
24
Incompl.
Record
A,2
No
27
Record
Data
Yes
No
A3/18
A3/02
A3/13

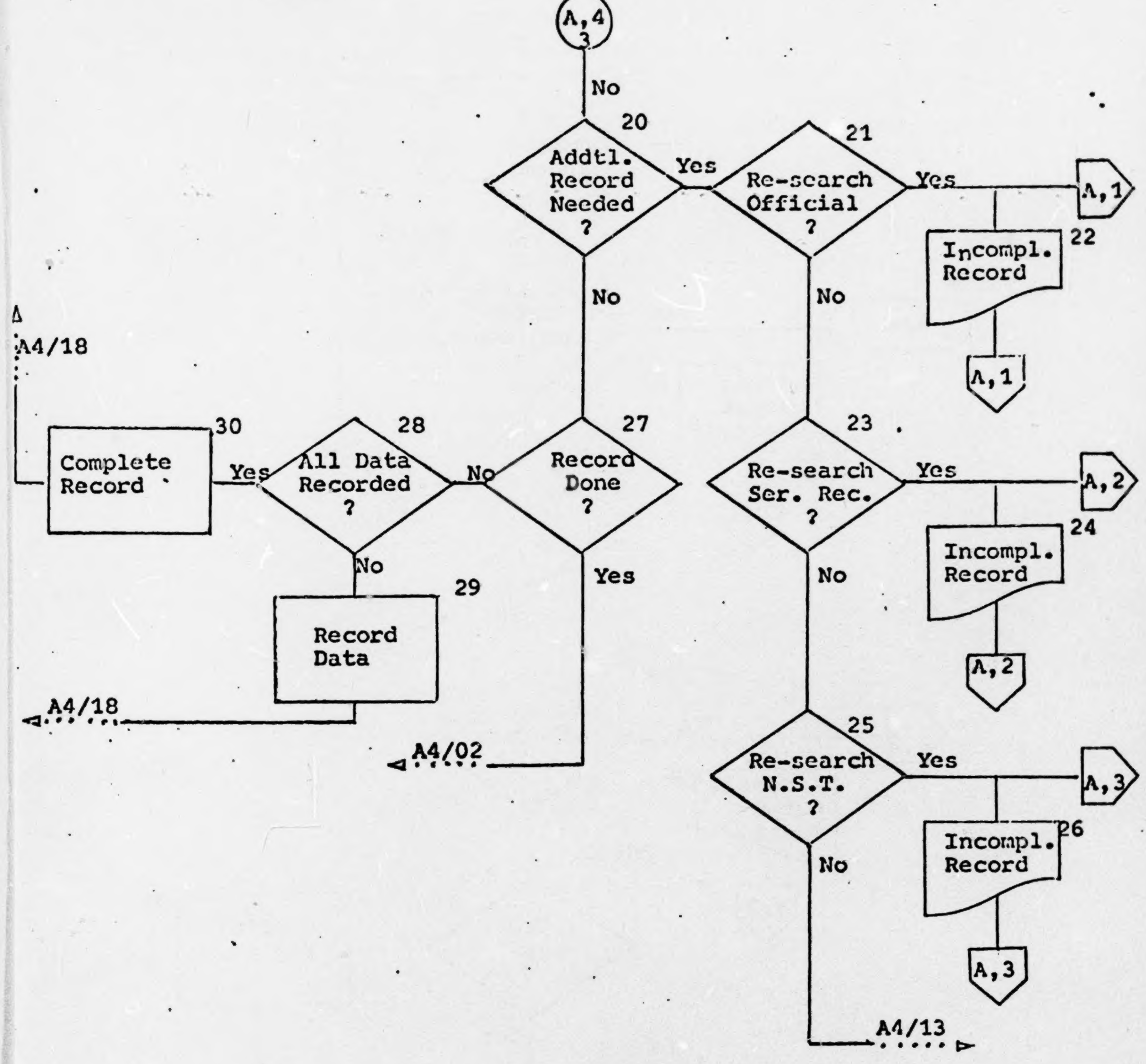
NSPP - Operations
A,4 - Verification
(ULS)


A,4
01
1x1 Output Sheet From A,3
A4/11
A4/27
02
All Searched In U.L.S.
Yes
03
Found ?
No
A,1
Incompl. Records
B,1
Yes
No
06
Last Record ?
Yes
07
Incompl. Record
B,1
05
Compl. Record
B,1
No
08
Pre 1960 Imprint ?
No
A,3
09
Non-ULS Record
A,3
Yes
10
Search Next Record
A,4 2

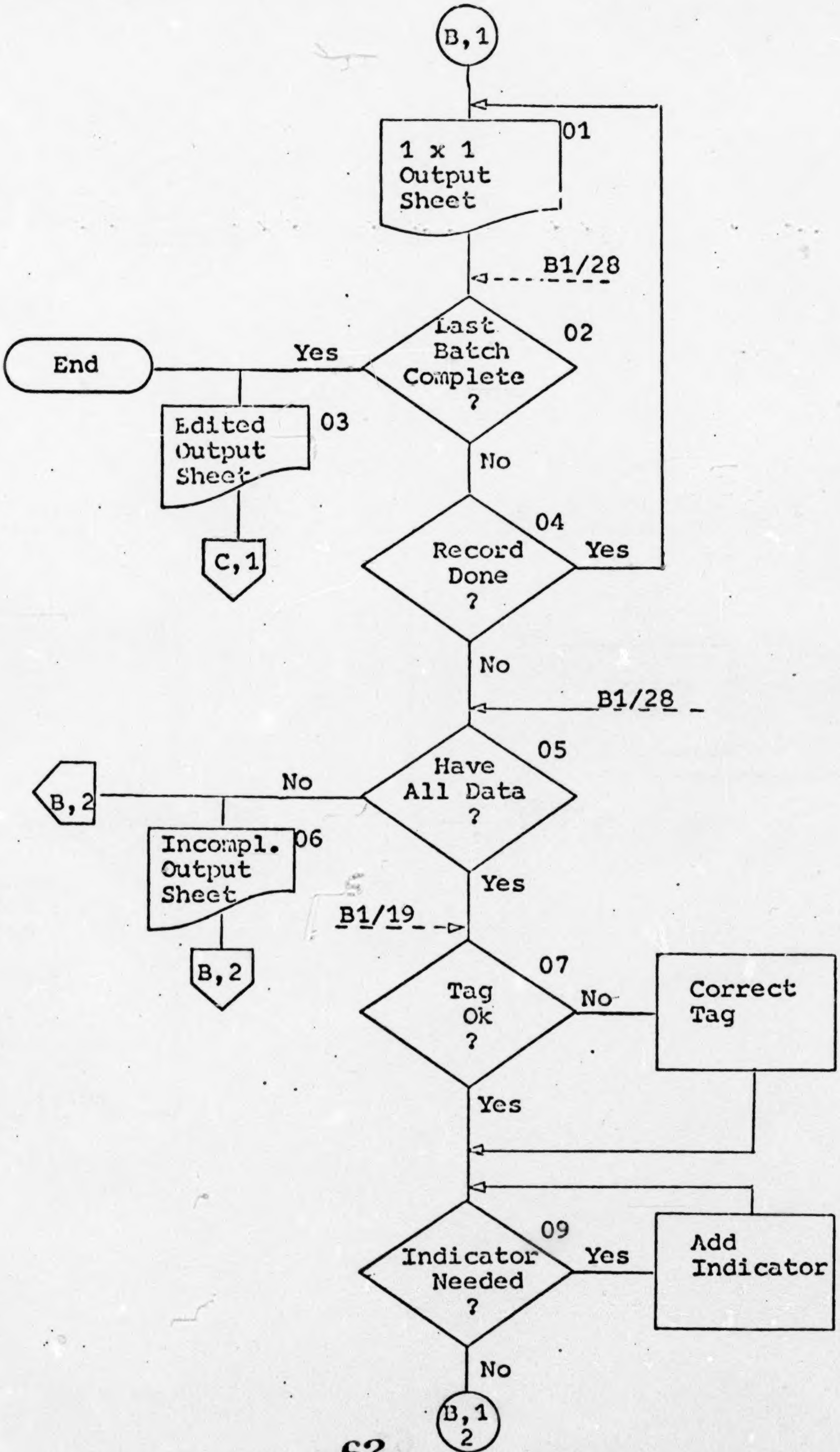
NSPP - Operations
A,4.2 - Verification
(ULS)

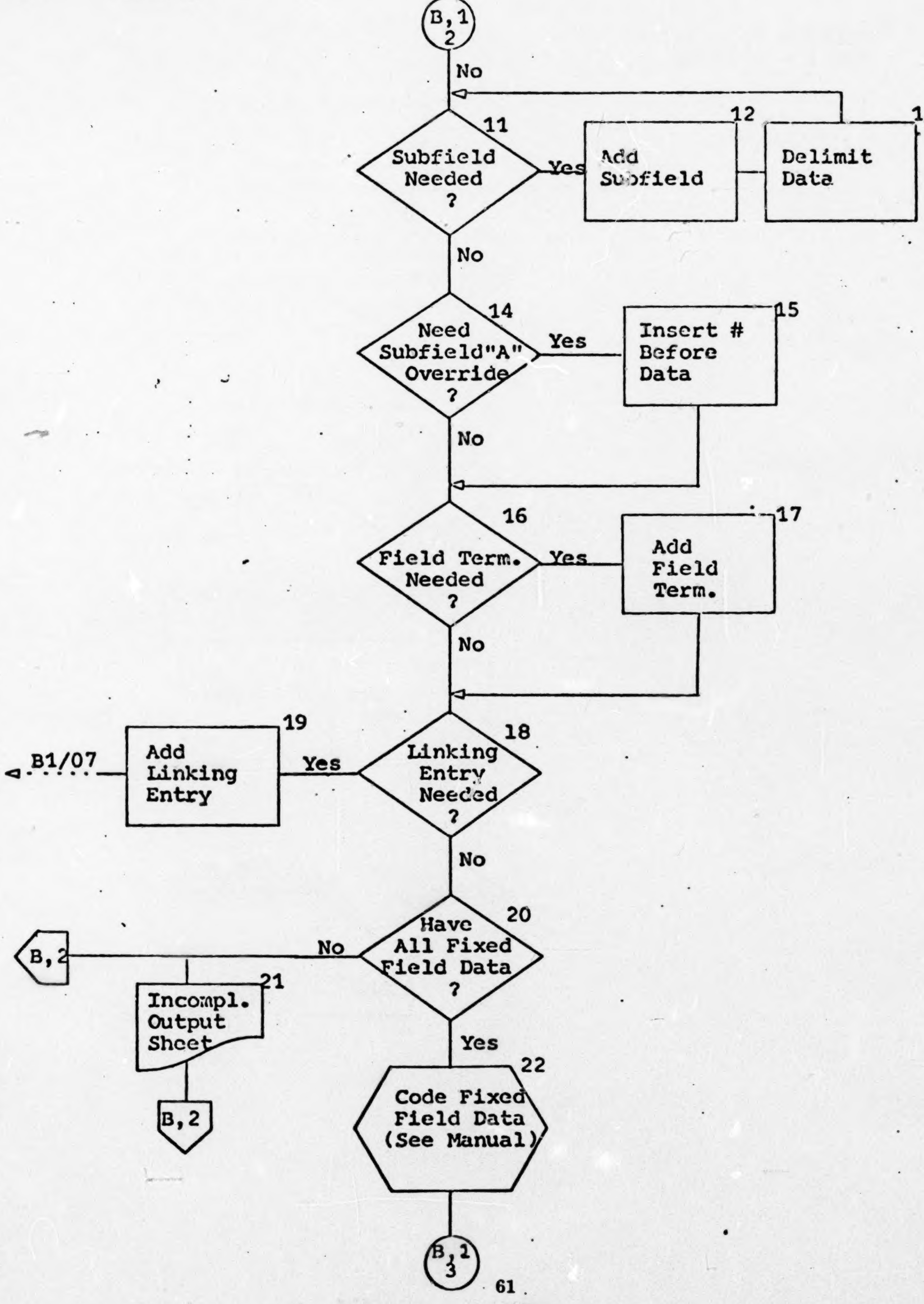

A,4 2
11
In ULS ?
No
A4/02
12
Univeri. Record
Yes
A4/25
13
Variant Form ?
Yes
Note Variation(s)
No
15
Re-search Official ?
Yes
A,
16
Incompl. Record
A,1
No
Compare Each Field Char.-by-Char. & Correct
A4/29
A4/30
19
Indicate Variant Tag
Yes
18
Variant Tag ?
No
A,4 3

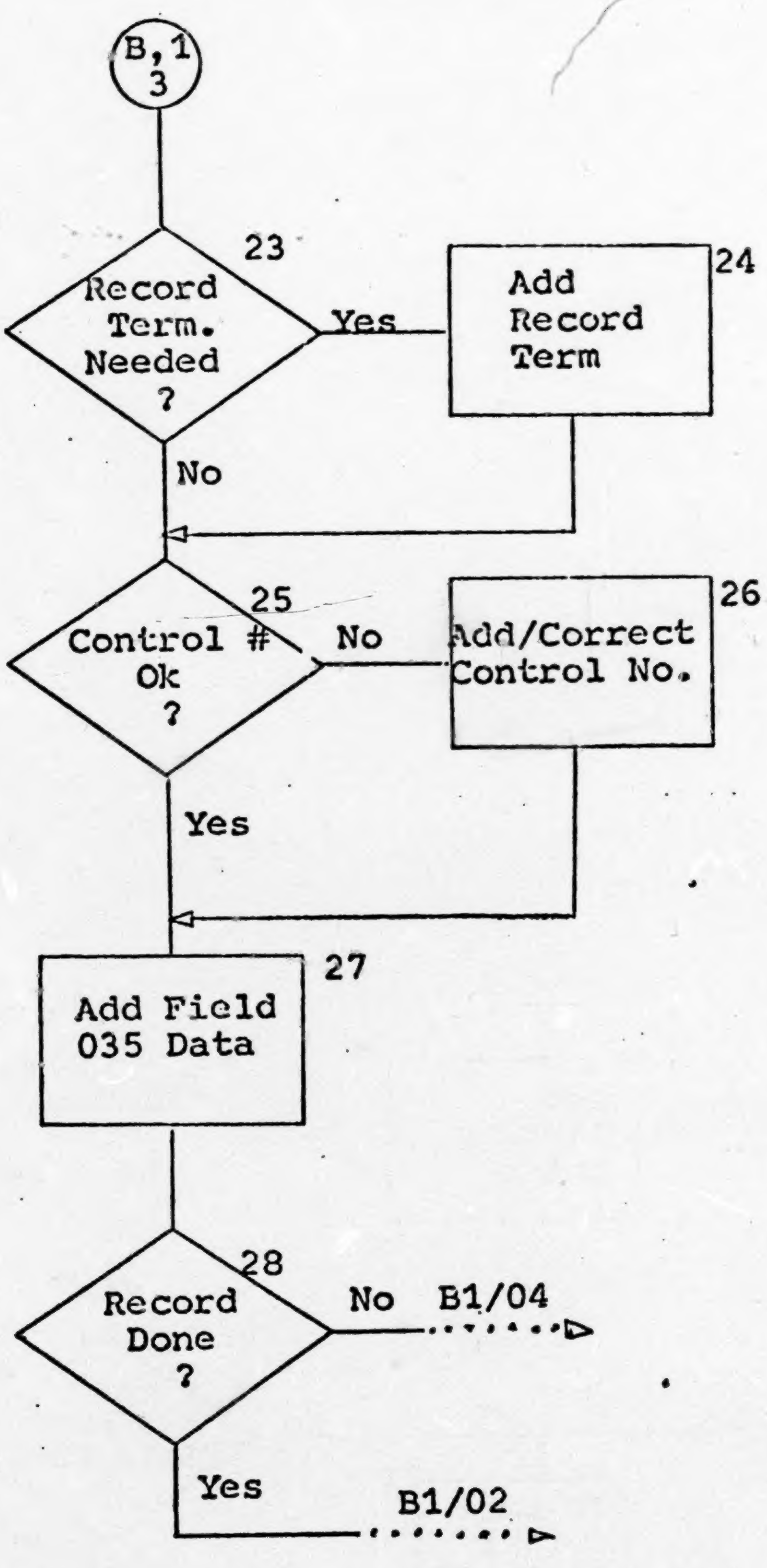
NSPP - Operations
A,4.3 - Verification
(ULS)


A,4
3
No
20
Addtl. Record Needed ?
Yes
21
Re-search Official ?
Yes
A,1
22
Incompl. Record
A,1
No
No
A4/18
30
Complete Record
Yes
28
All Data Recorded ?
No
27
Record Done ?
23
Re-search Ser. Rec. ?
Yes
A,2
24
Incompl. Record
A,2
No
Yes
No
29
Record Data
A4/18
A4/02
25
Re-search N.S.T. ?
Yes
A,3
26
Incompl. Record
A,3
No
A4/13

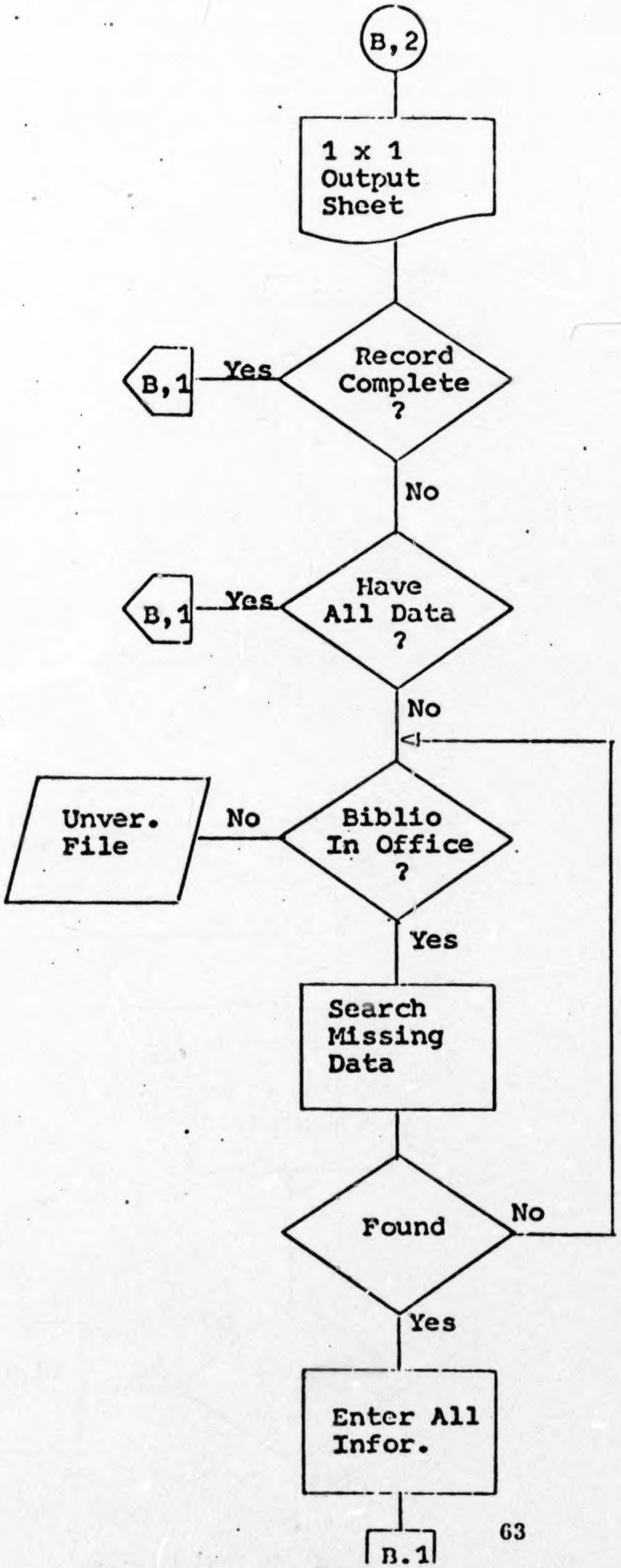
NSPP - Operations
B,1 - Editing


B,1
1 x 1 Output Sheet
01
B1/28
Last Batch Complete ?
02
Yes
End
Edited Output Sheet
03
C,1
No
Record Done ?
04
Yes
No
B1/28
Have All Data ?
05
No
B,2
Incompl. Output Sheet
06
B,2
Yes
B1/19
Tag Ok ?
07
No
Correct Tag
08
Yes
Indicator Needed ?
09
Yes
Add Indicator
10
No
B,1
2

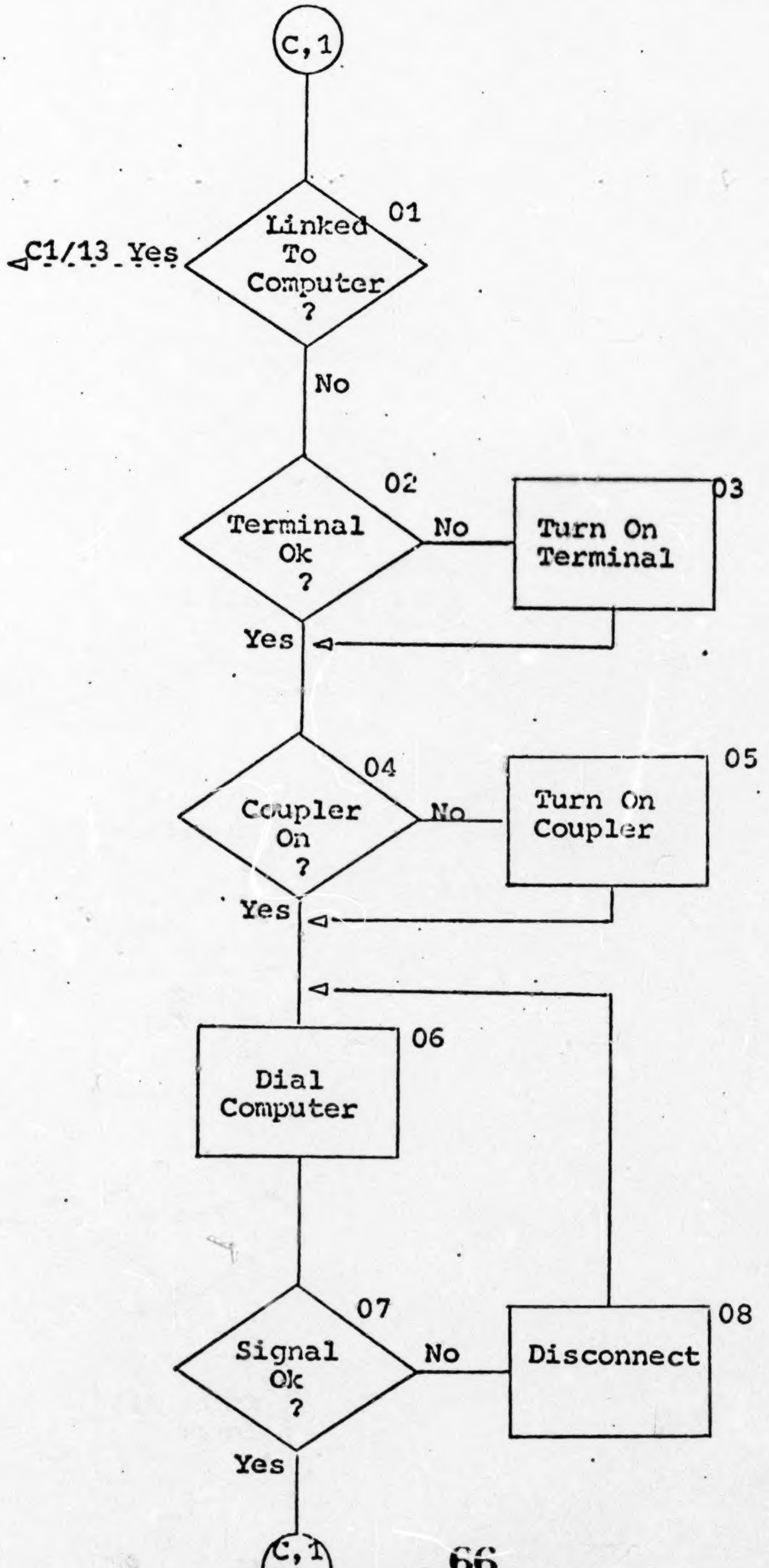

B,1
2
No
11
Subfield Needed ?
Yes
12
Add Subfield
13
Delimit Data
No
14
Need Subfield"A" Override ?
Yes
15
Insert # Before Data
No
16
Field Term. Needed ?
Yes
17
Add Field Term.
No
19
Add Linking Entry
B1/07
Yes
18
Linking Entry Needed ?
No
20
Have All Fixed Field Data ?
No
B,2
21
Incompl. Output Sheet
B,2
Yes
22
Code Fixed Field Data (See Manual)
B,1
3


B,1
3
23
Record
Term.
Needed
?
Yes
24
Add
Record
Term
No
25
Control #
Ok
?
No
26
Add/Correct
Control No.
Yes
27
Add Field
035 Data
28
Record
Done
?
No
B1/04
Yes
B1/02

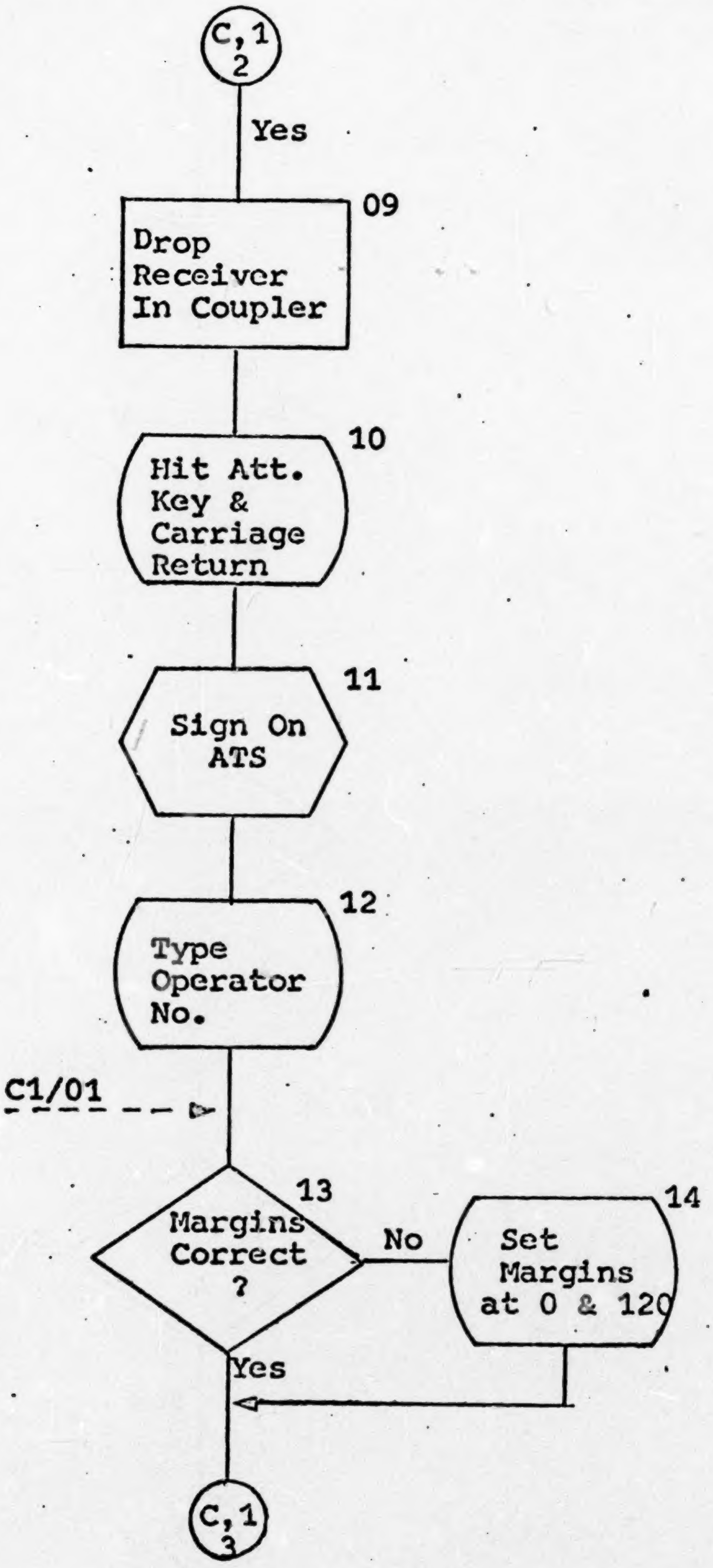
NSPP - Operations
B,2 - Editing - Supplementary Search


B,2
1 x 1 Output Sheet
Record Complete ?
Yes
B,1
No
Have All Data ?
Yes
B,1
No
Biblio In Office ?
No
Unver. File
Yes
Search Missing Data
Found
No
Yes
Enter All Infor.
B.1

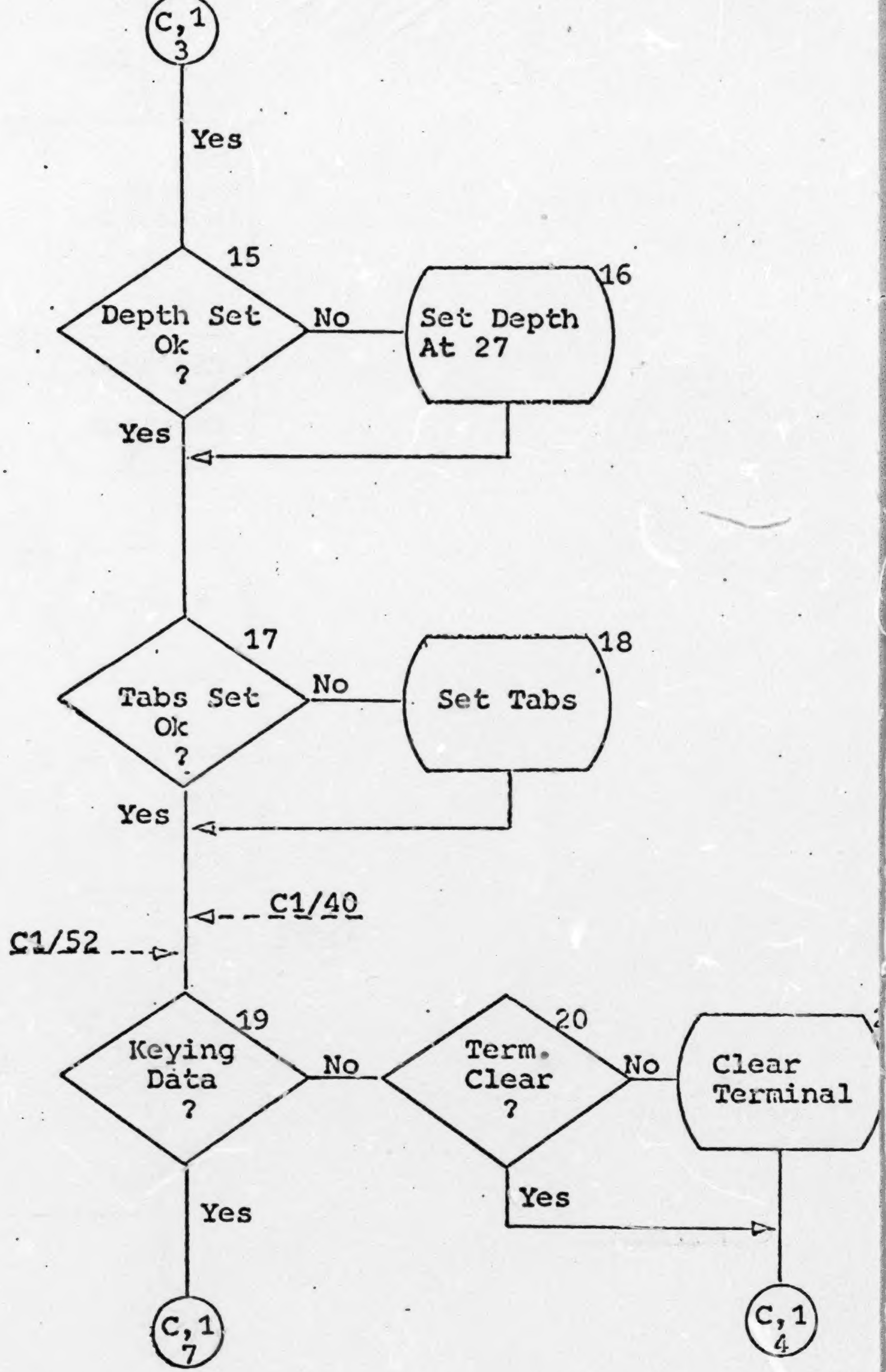
NSPP - Operations
C,1 - Key Input

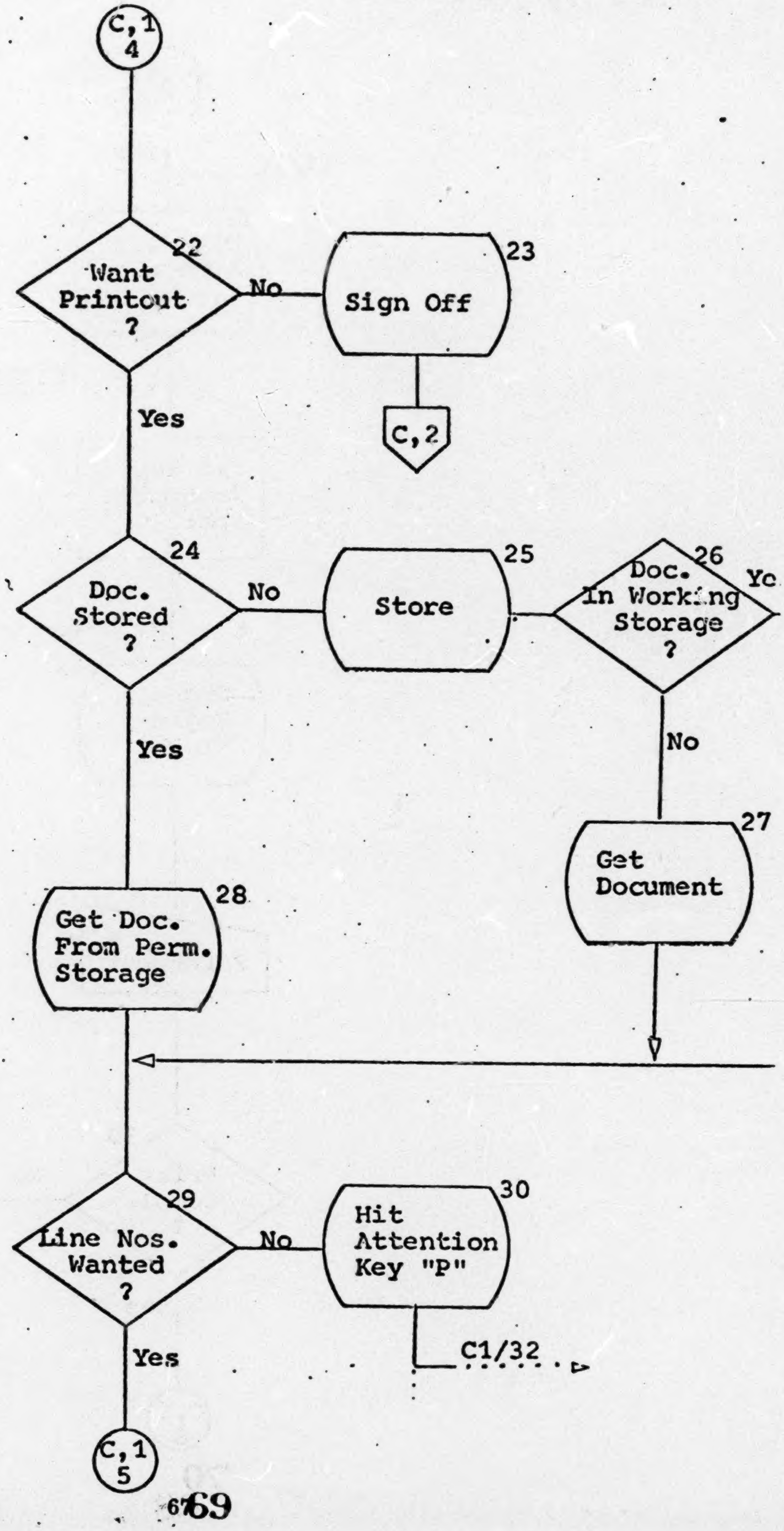

C,1
01
Linked To Computer ?
C1/13 Yes
No
02
Terminal Ok ?
No
03
Turn On Terminal
Yes
04
Coupler On ?
No
05
Turn On Coupler
Yes
06
Dial Computer
07
Signal Ok ?
No
08
Disconnect
Yes
C,1 2

NSPP - Operations
C,1.2 - Key Input

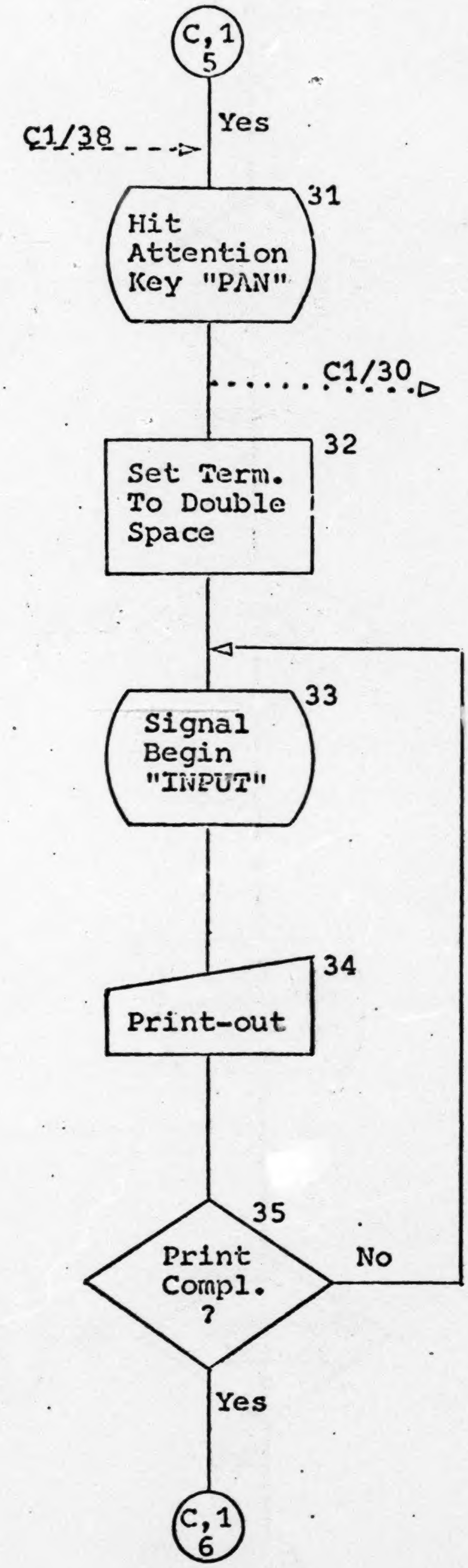

C,1
2
Yes
09
Drop
Receiver
In Coupler
10
Hit Att.
Key &
Carriage
Return
11
Sign On
ATS
12
Type
Operator
No.
C1/01
13
Margins
Correct
?
No
14
Set
Margins
at 0 & 120
Yes
C,1
3

NSPP - Operations
C,1.3 - Key Input

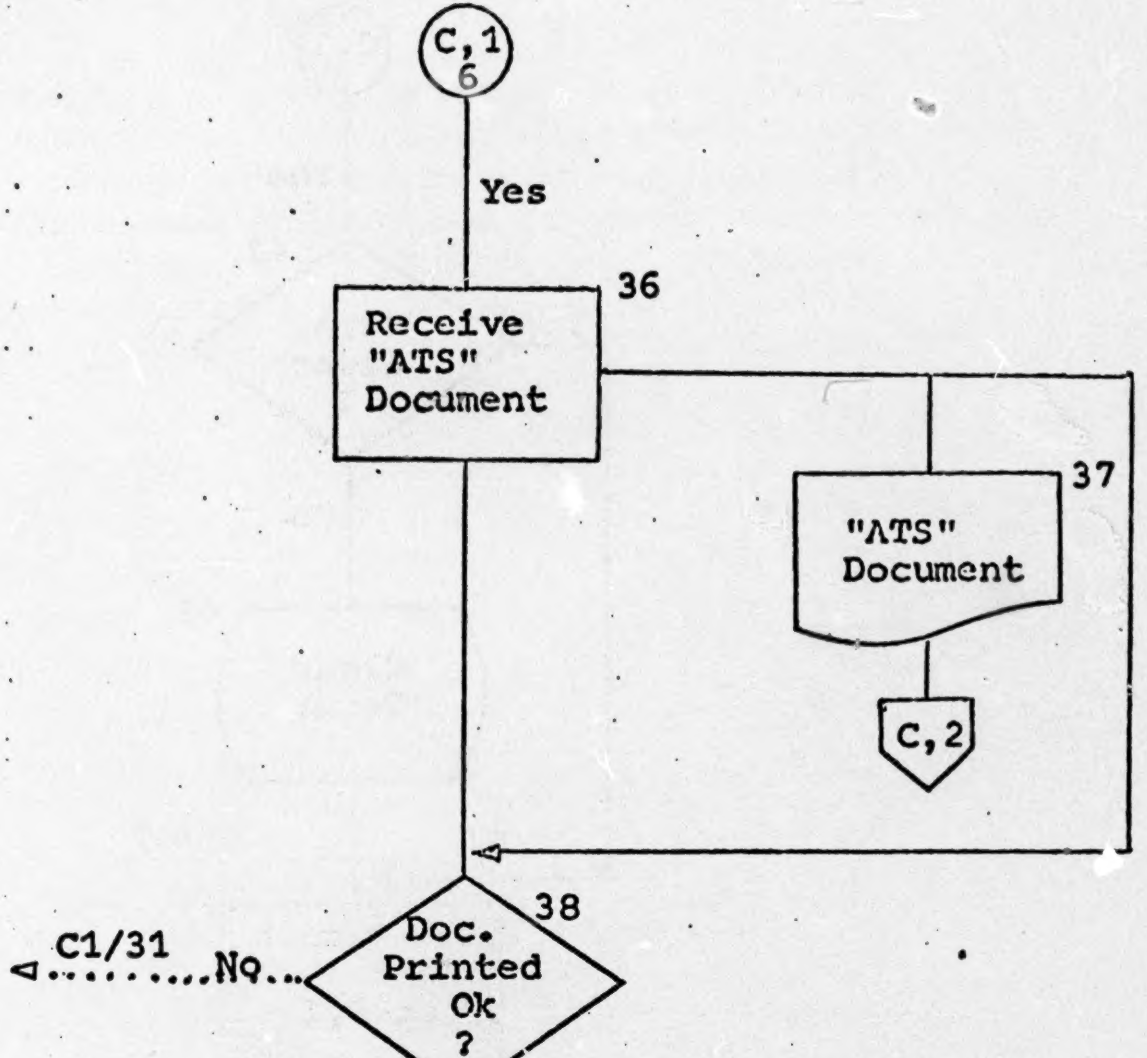

C,1
3
Yes
15
Depth Set
Ok
?
No
16
Set Depth
At 27
Yes
17
Tabs Set
Ok
?
No
18
Set Tabs
Yes
C1/40
C1/52
19
Keying
Data
?
No
20
Term.
Clear
?
No
Clear
Terminal
Yes
Yes
C,1
7
C,1
4


C,1
4
22
Want Printout ?
No
23
Sign Off
C,2
Yes
24
Doc. Stored ?
No
25
Store
26
Doc. In Working Storage ?
Yc
No
27
Get Document
Yes
28
Get Doc. From Perm. Storage
29
Line Nos. Wanted ?
No
30
Hit Attention Key "P"
C1/32
Yes
C,1
5

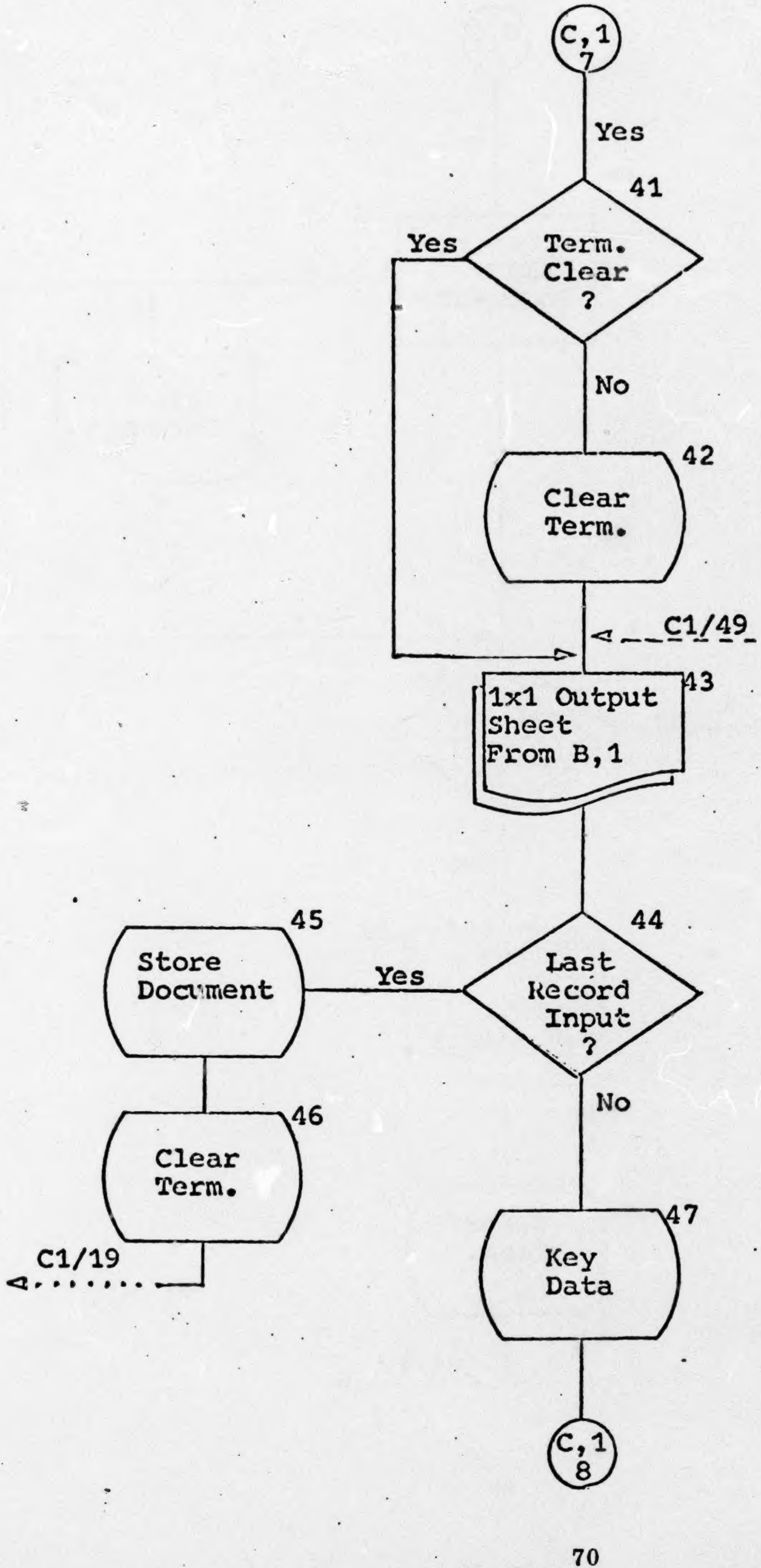
NSPP - Operations
C,1.5 - Key Input


C,1
5
Yes
C1/38
31
Hit
Attention
Key "PAN"
C1/30
32
Set Term.
To Double
Space
33
Signal
Begin
"INPUT"
34
Print-out
35
Print
Compl.
?
No
Yes
C,1
6

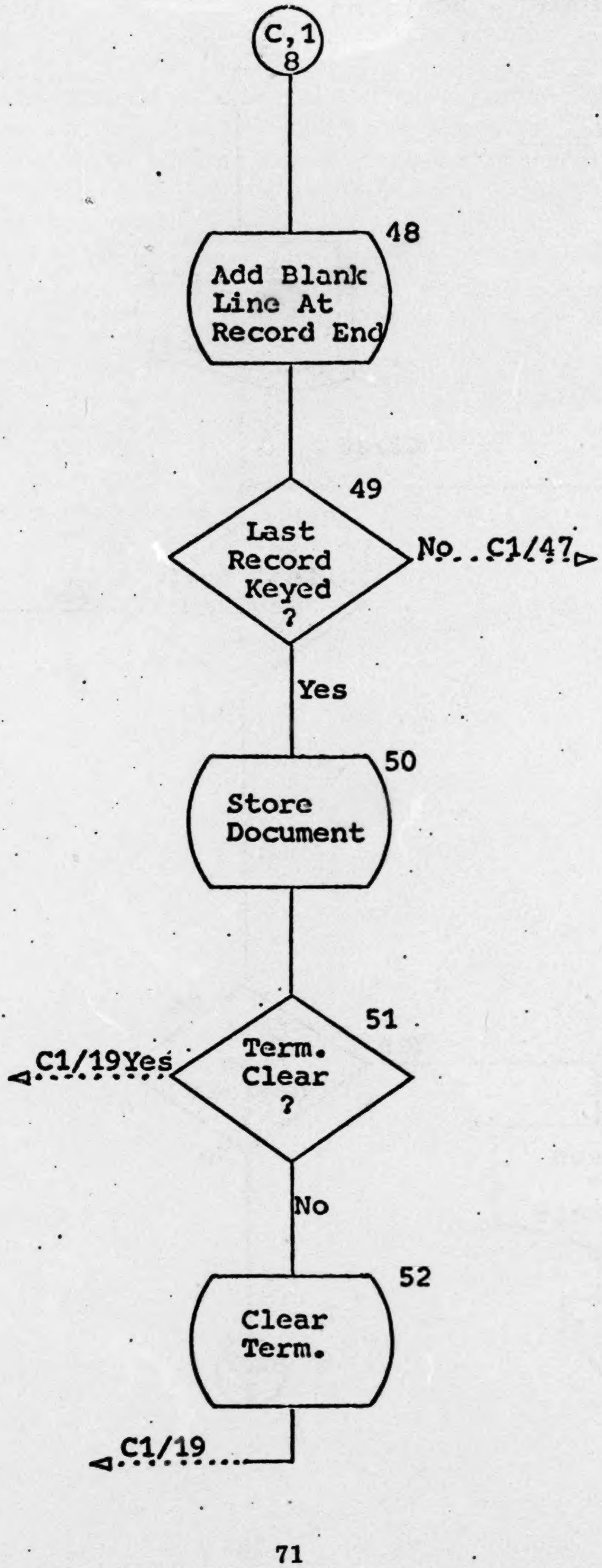
NSPP - Operations
C,1.6 - Key Input


C,1 6
Yes
36
Receive "ATS" Document
37
"ATS" Document
C,2
38
Doc. Printed Ok ?
C1/31
No
Yes
39
Collate With Orginal Work Sheet
40
Clear Term.
C1/19

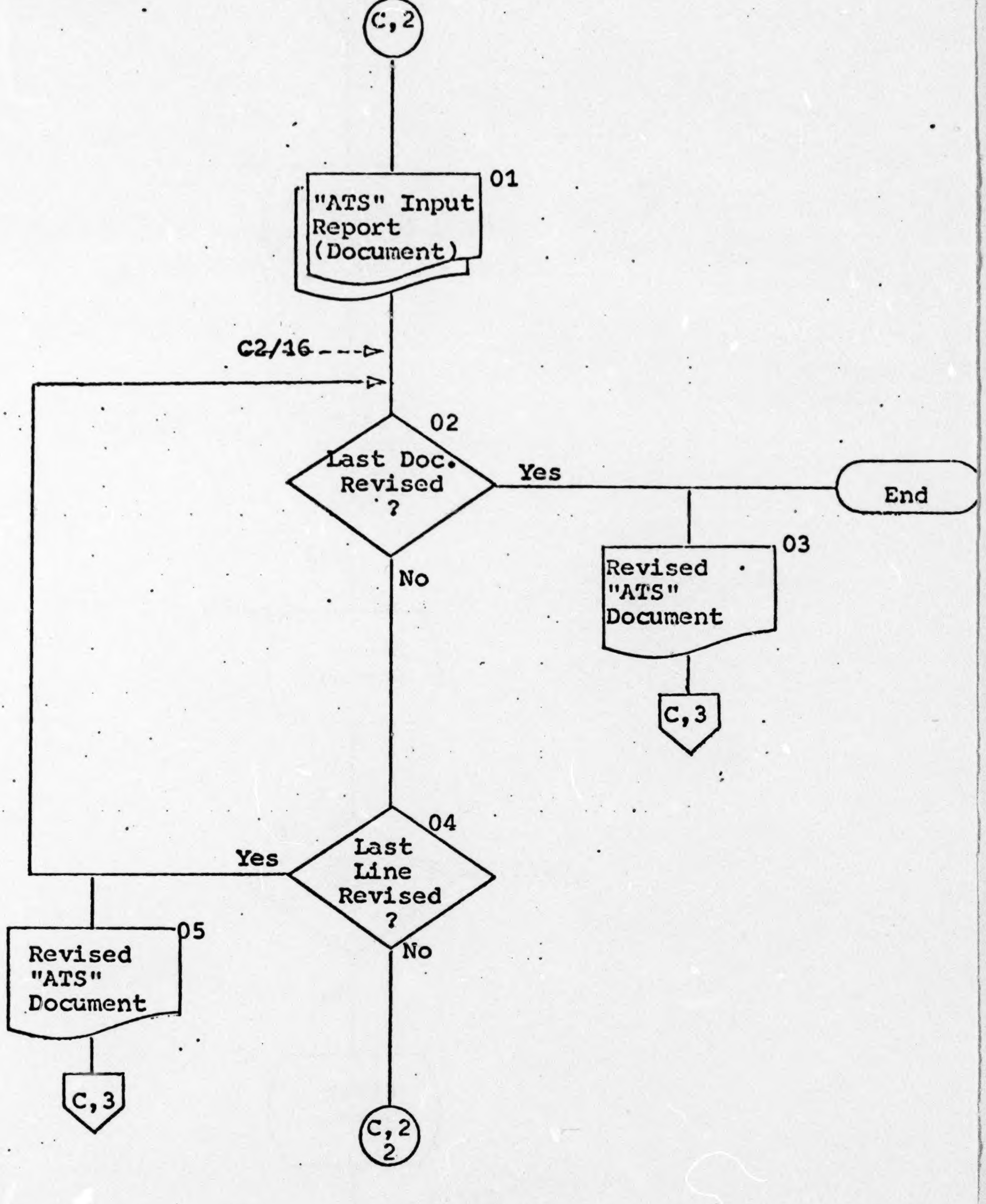
NSPP - Operations
C,1.7 - Key Input

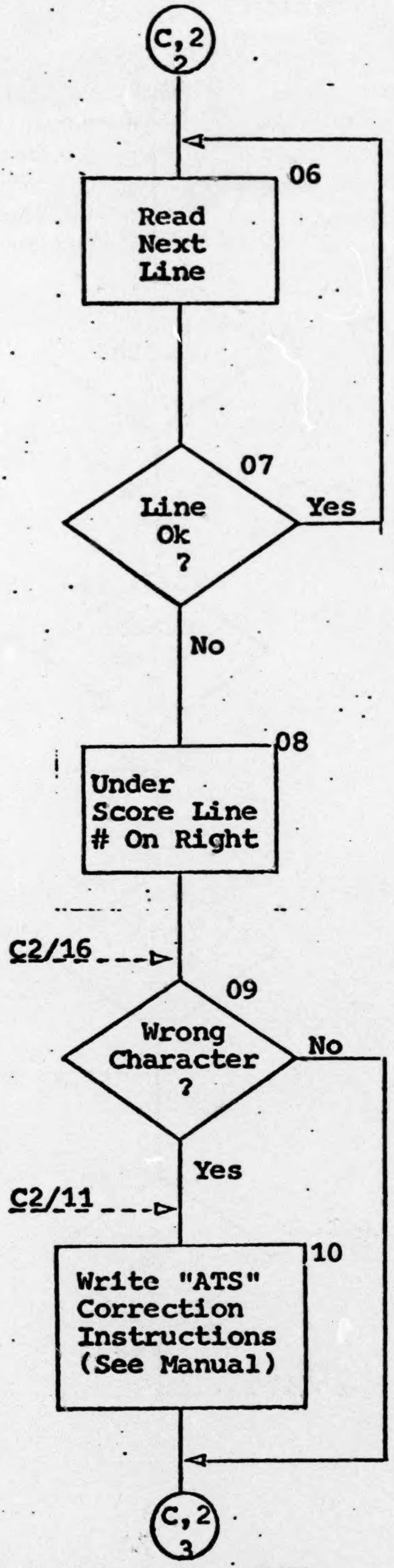

C,1
7
Yes
41
Term.
Clear
?
Yes
No
42
Clear
Term.
C1/49
1x1 Output
Sheet
From B,1
43
44
Last
Record
Input
?
Yes
45
Store
Document
46
Clear
Term.
C1/19
No
47
Key
Data
C,1
8

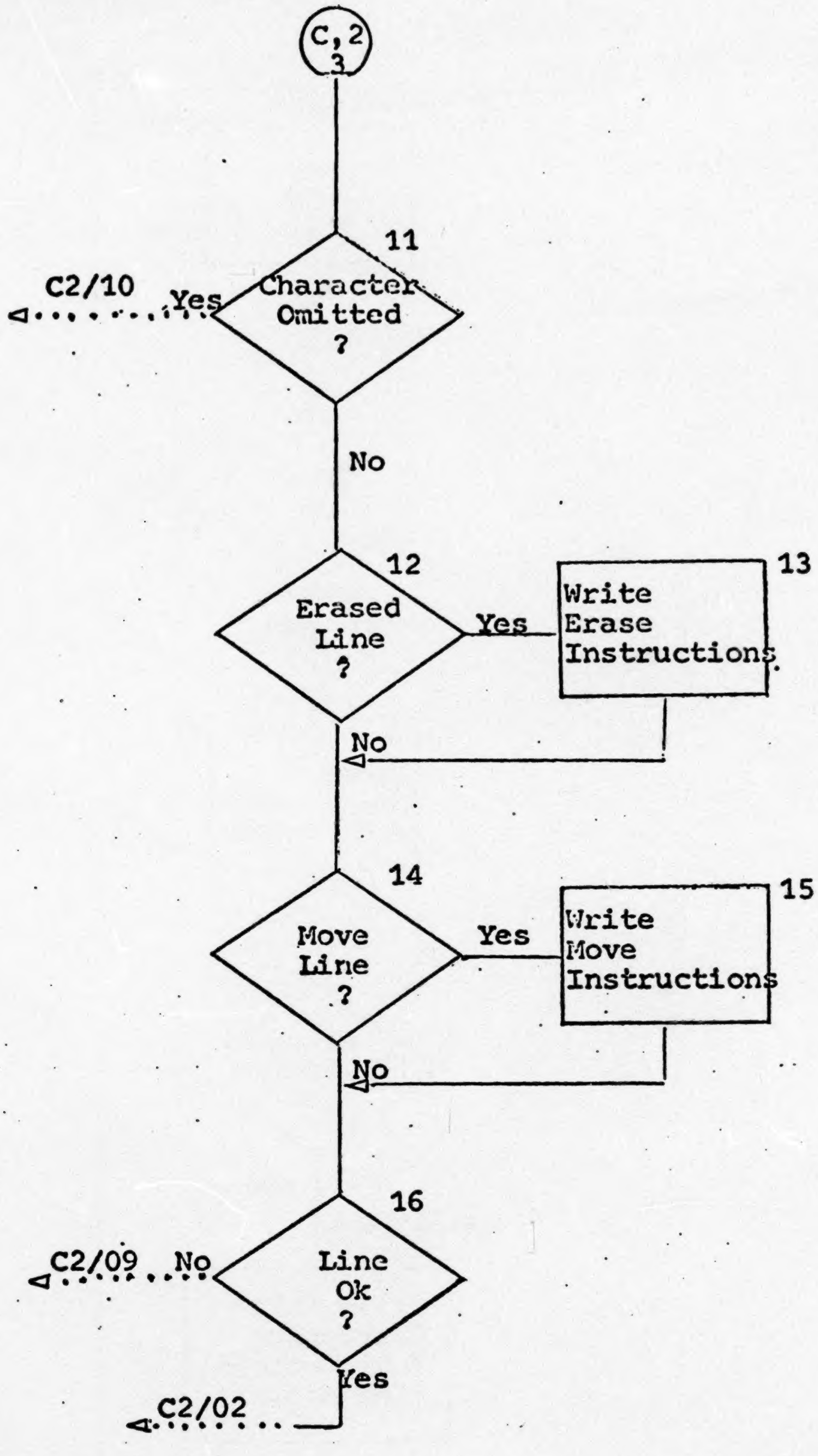
NSPP - Operations
C,1.8 - Key Input


C,1
8
48
Add Blank Line At Record End
49
Last Record Keyed ?
No..C1/47
Yes
50
Store Document
51
Term. Clear ?
C1/19Yes
No
52
Clear Term.
C1/19

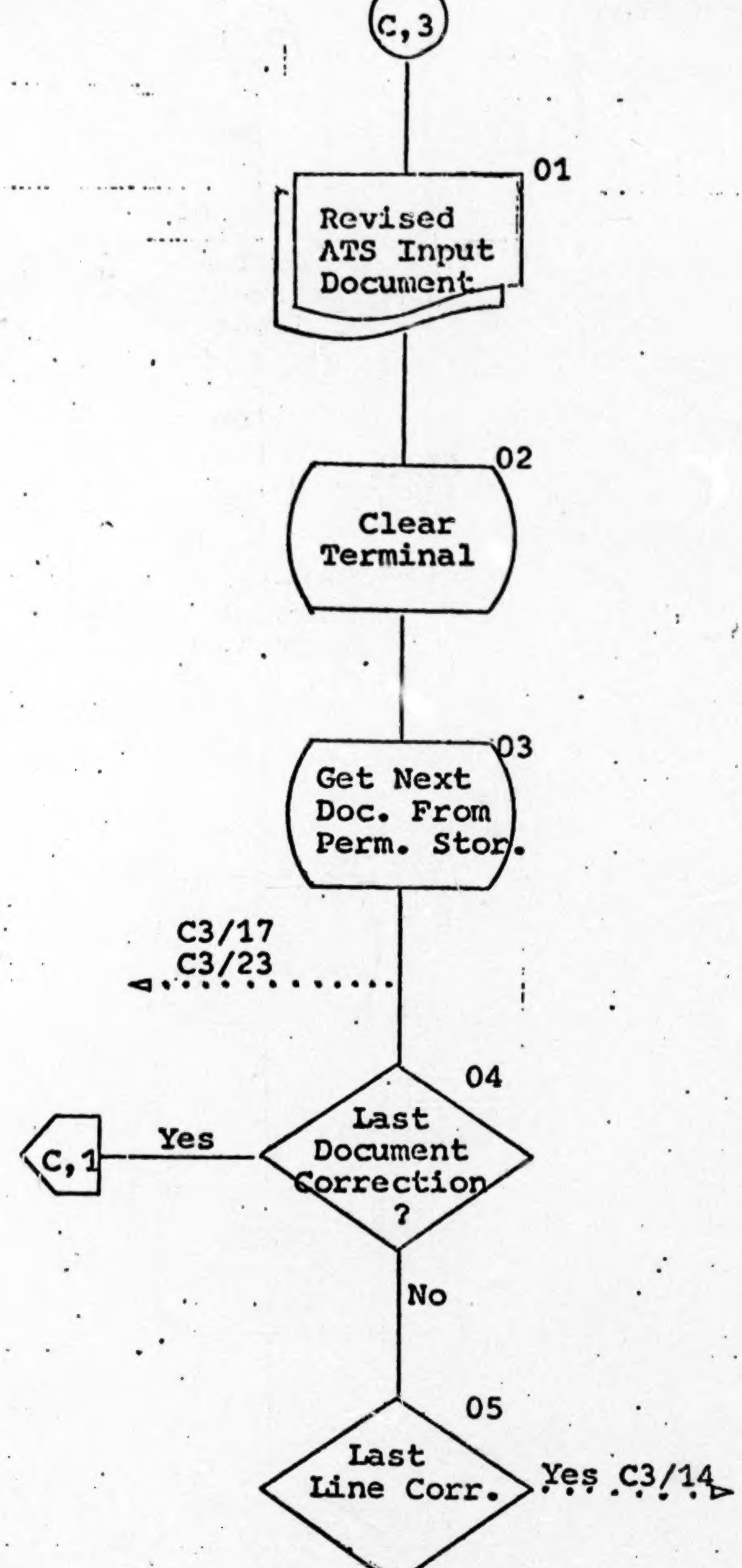
NSPP - Operations
C,2 - Key Input - Revision


C,2
01
"ATS" Input Report (Document)
C2/16
02
Last Doc. Revised ?
Yes
End
03
Revised "ATS" Document
C,3
No
04
Last Line Revised ?
Yes
05
Revised "ATS" Document
C,3
No
C,2
2

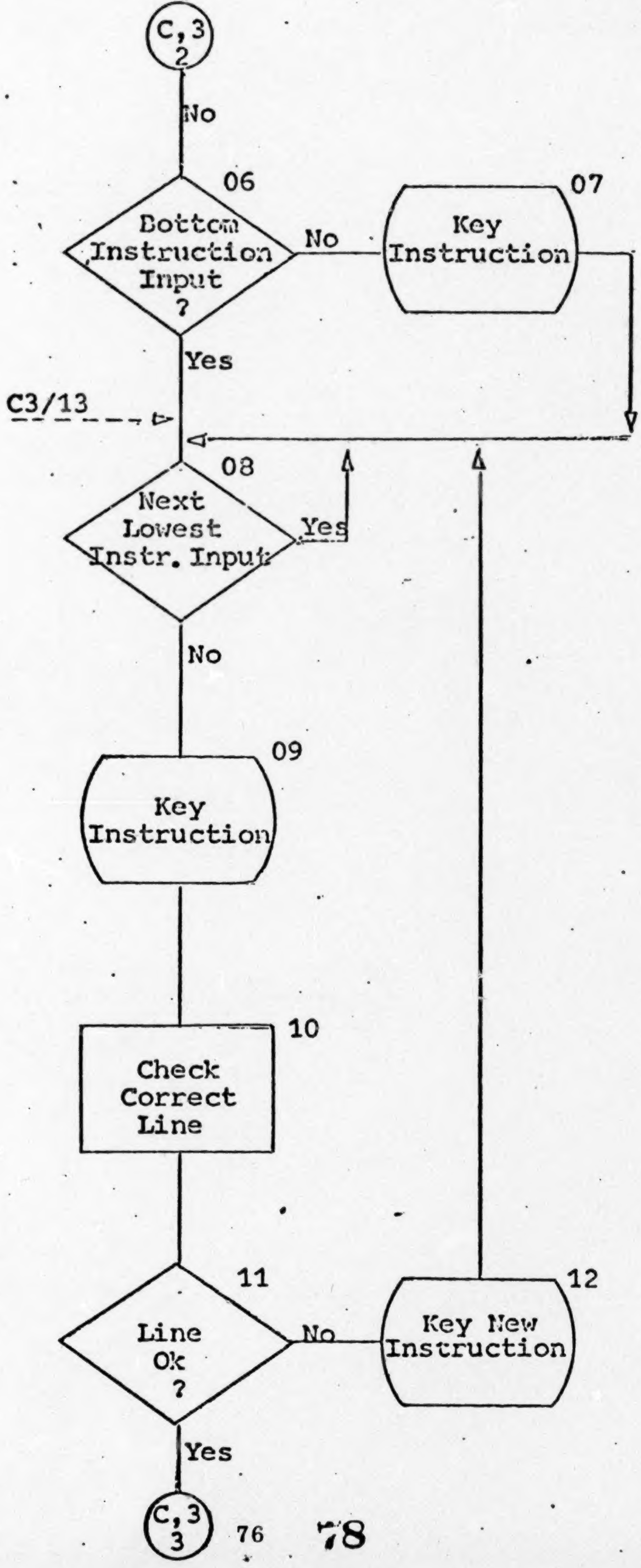

C,2
2
06
Read Next Line
07
Line Ok ?
Yes
No
08
Under Score Line # On Right
C2/16
09
Wrong Character ?
No
Yes
C2/11
10
Write "ATS" Correction Instructions (See Manual)
C,2
3


C,2
3
11
Character Omitted ?
Yes
C2/10
No
12
Erased Line ?
Yes
13
Write Erase Instructions
No
14
Move Line ?
Yes
15
Write Move Instructions
No
16
Line Ok ?
No
C2/09
Yes
C2/02

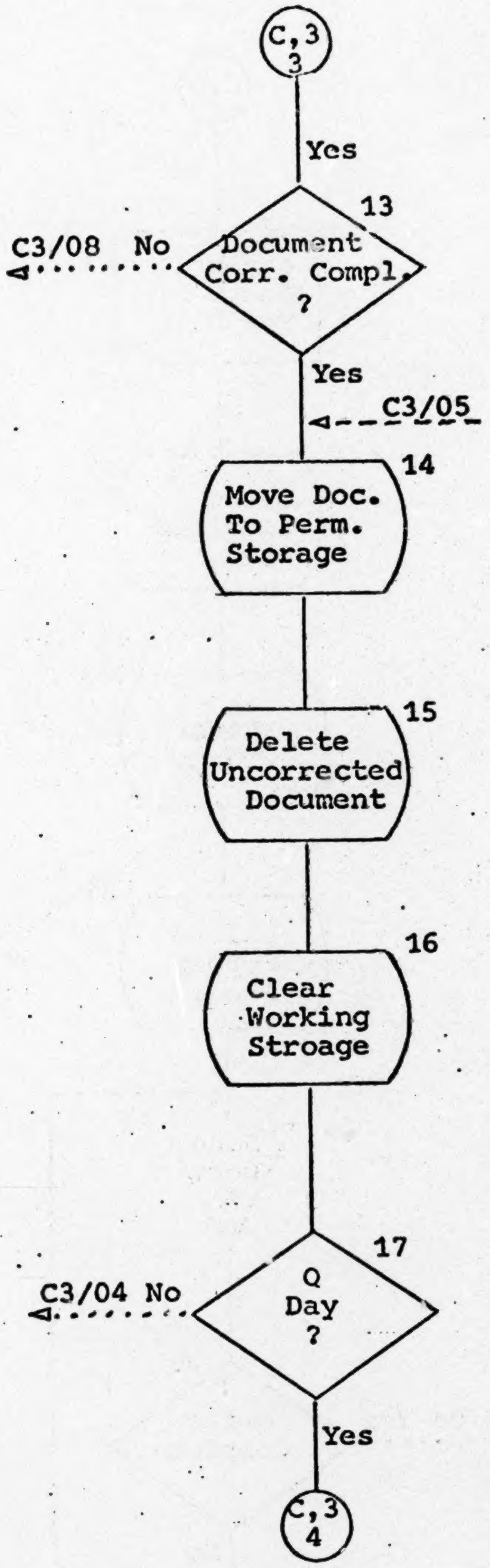
NSPP - Operations
C,3 - Key Input - Correction


C,3
01
Revised ATS Input Document
02
Clear Terminal
03
Get Next Doc. From Perm. Stor.
C3/17
C3/23
04
Last Document Correction ?
Yes
C,1
No
05
Last Line Corr.
Yes C3/14
No
C,3 2

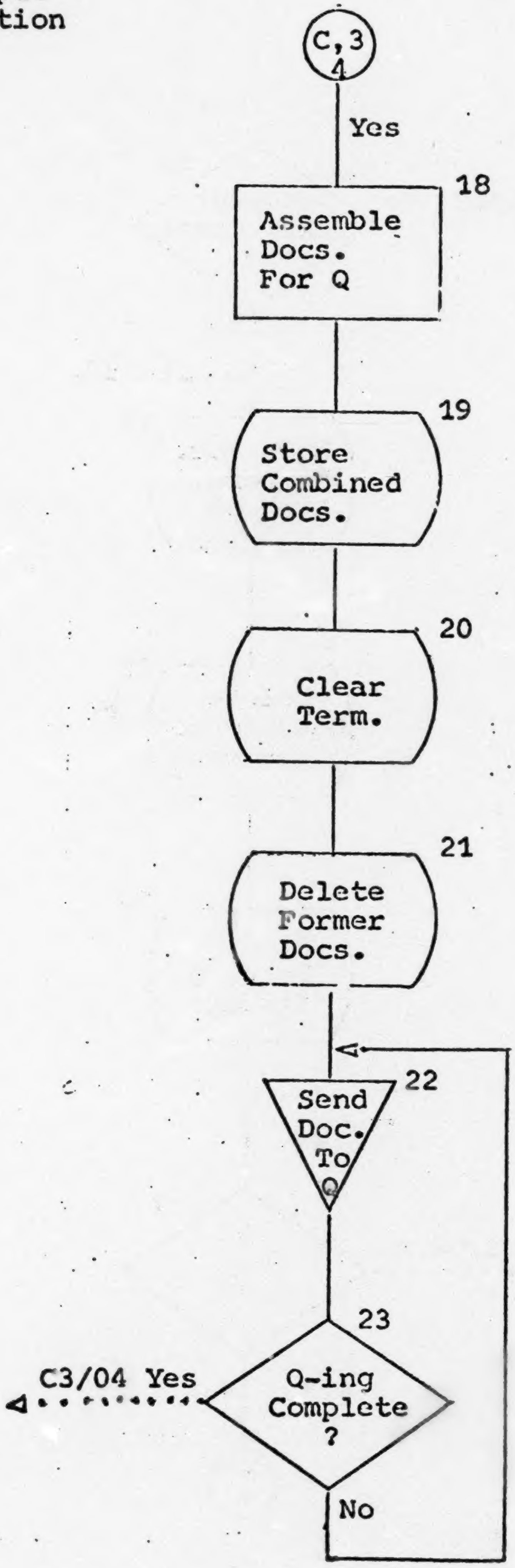
NSPP - Operations
C,3.2 - Key Input -
Correction


C,3
2
No
06
Bottom
Instruction
Input
?
No
07
Key
Instruction
Yes
C3/13
08
Next
Lowest
Instr. Input
Yes
No
09
Key
Instruction
10
Check
Correct
Line
11
Line
Ok
?
No
12
Key New
Instruction
Yes
C,3
3

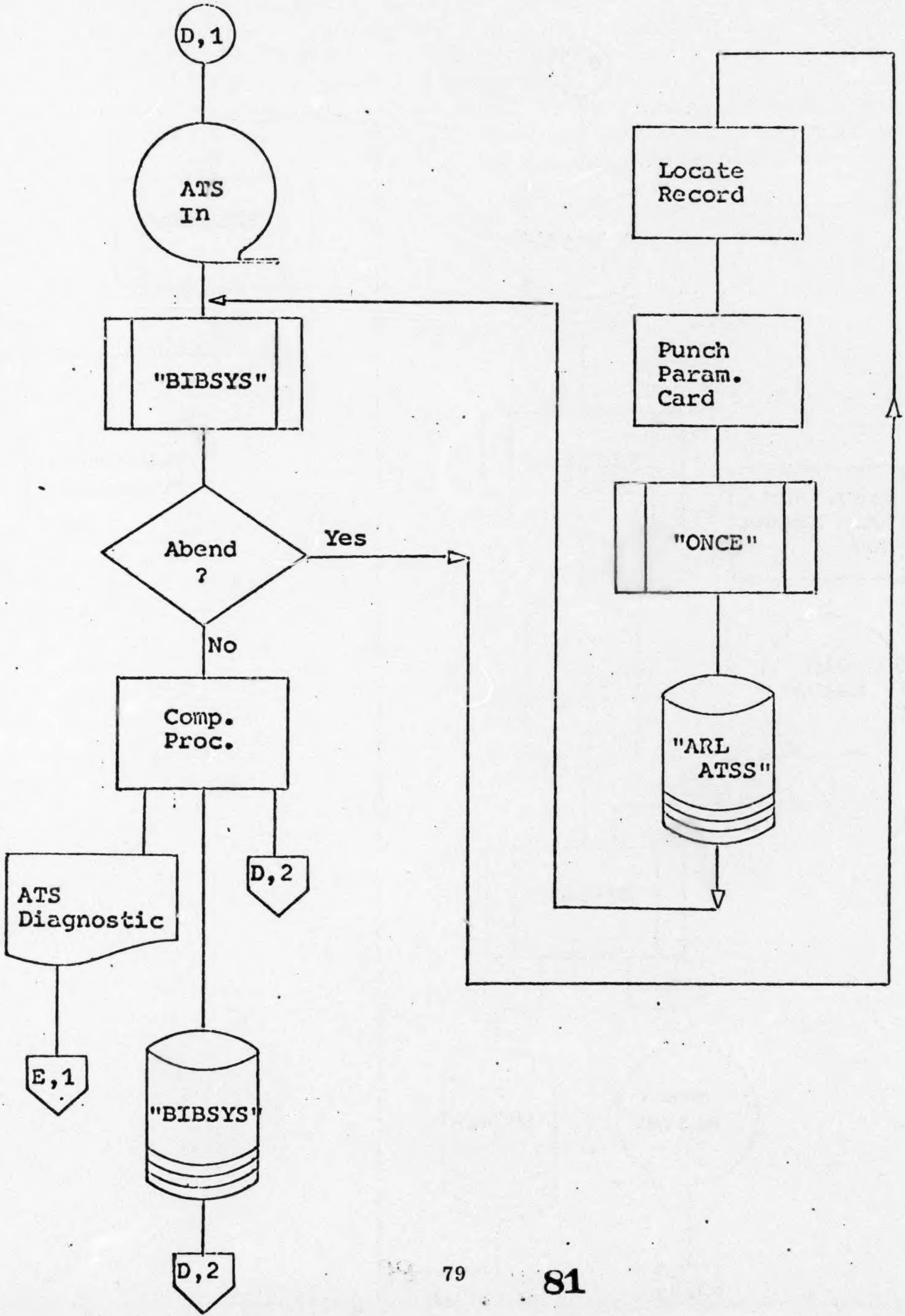
NSPP - Operations
C,3.3 - Key Input -
Correction


C,3
3
Yes
13
C3/08 No
Document
Corr. Compl.
?
Yes
C3/05
14
Move Doc.
To Perm.
Storage
15
Delete
Uncorrected
Document
16
Clear
Working
Stroage
17
C3/04 No
Q
Day
?
Yes
C,3
4

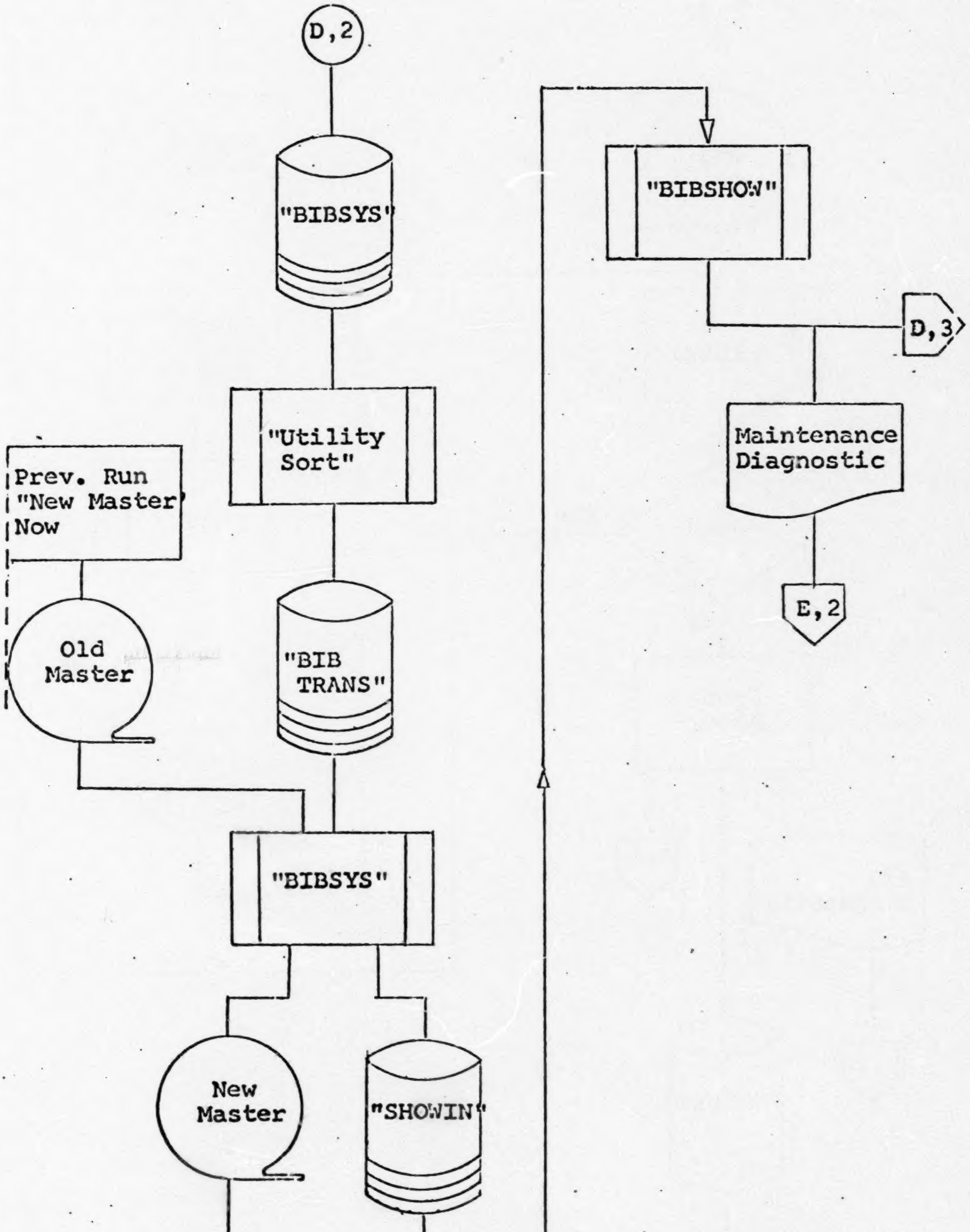
NSPP - Operations
C,3.4 - Key Input -
Correction


C,3
4
Yes
18
Assemble
Docs.
For Q
19
Store
Combined
Docs.
20
Clear
Term.
21
Delete
Former
Docs.
22
Send
Doc.
To
Q
23
C3/04 Yes
Q-ing
Complete
?
No

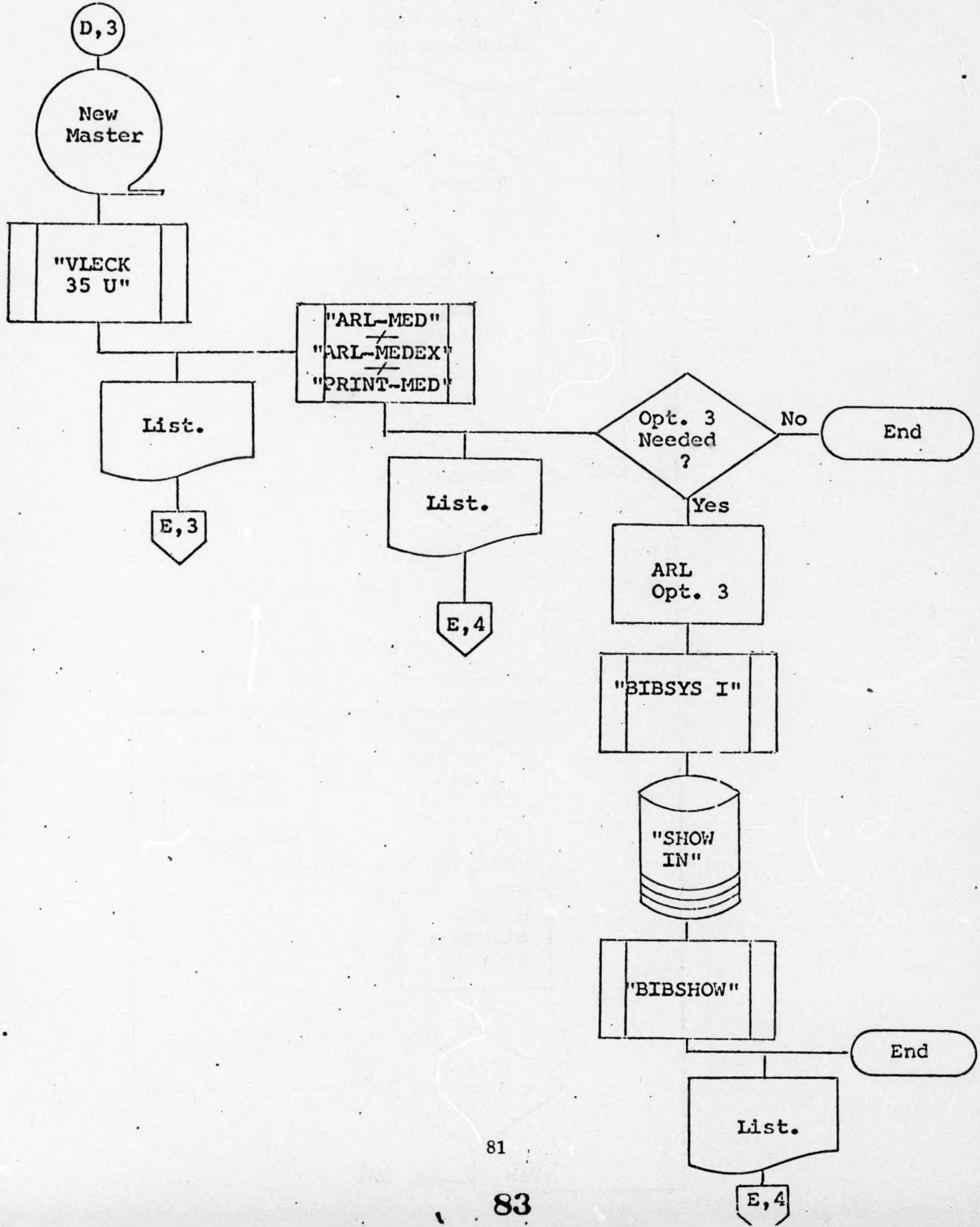
NSPP - Operations
D,1 - Update - run
"ATS"


D,1
ATS In
"BIBSYS"
Abend ?
Yes
No
Comp. Proc.
ATS Diagnostic
D,2
E,1
"BIBSYS"
D,2
Locate Record
Punch Param. Card
"ONCE"
"ARL ATSS"

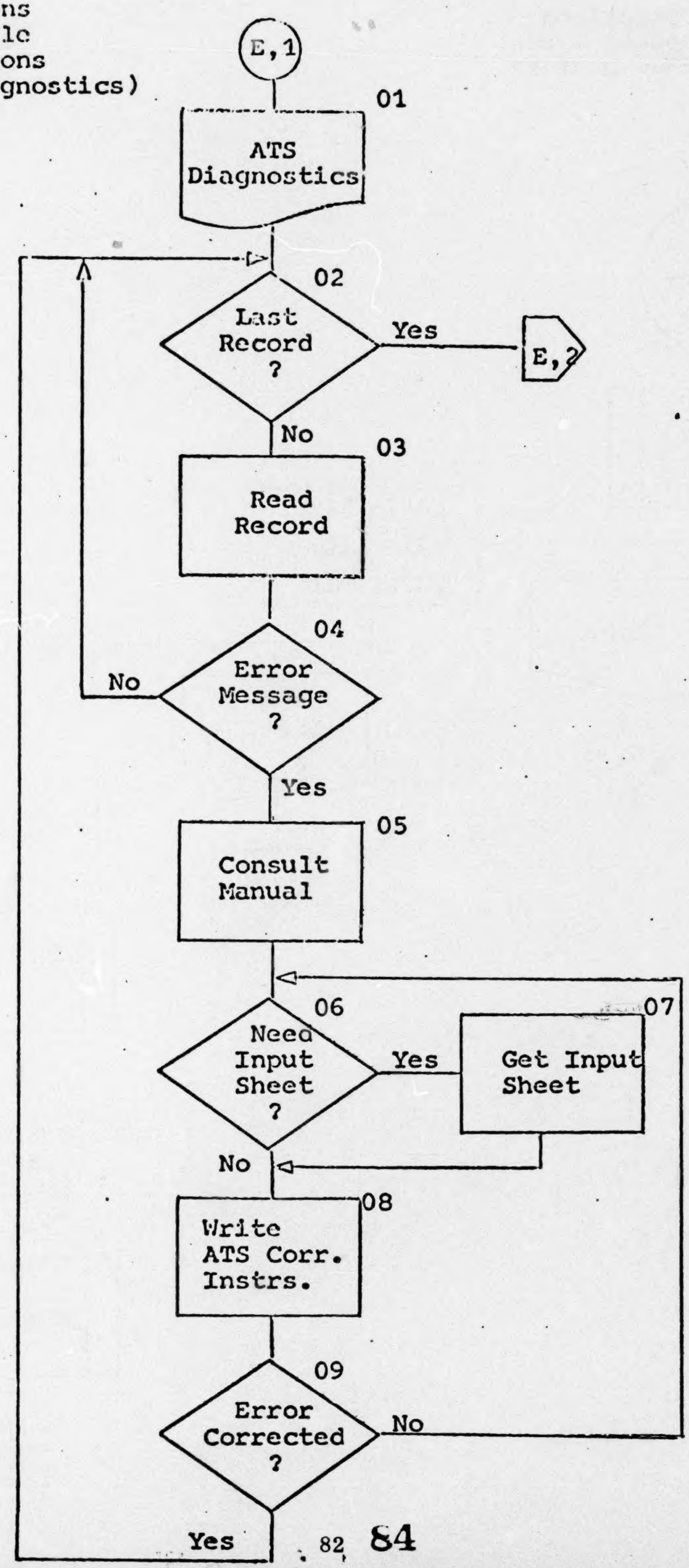
NSPP - Operations
D,2 - Update - run
"BIBSYS"


D,2
"BIBSYS"
"Utility Sort"
Prev. Run "New Master" Now
Old Master
"BIB TRANS"
"BIBSYS"
New Master
"SHOWIN"
D,3
"BIBSHOW"
D,3
Maintenance Diagnostic
E,2

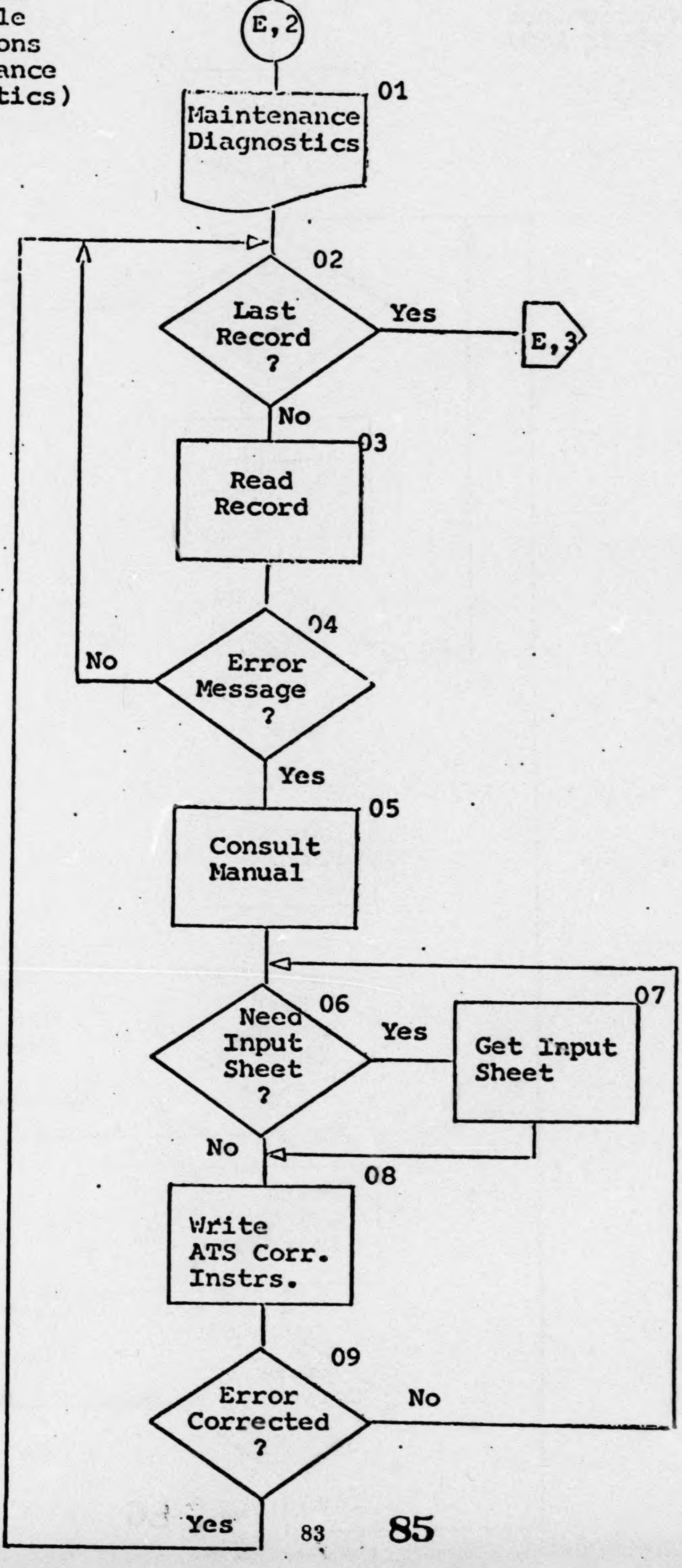
NSPP - Operations
D,3 - Update - run
"New Master"


D,3
New Master
"VLECK 35 U"
"ARL-MED" / "ARL-MEDEX" / "PRINT-MED"
List.
E,3
List.
E,4
Opt. 3 Needed ?
No
End
Yes
ARL Opt. 3
"BIBSYS I"
"SHOW IN"
"BIBSHOW"
End
List.
E,4

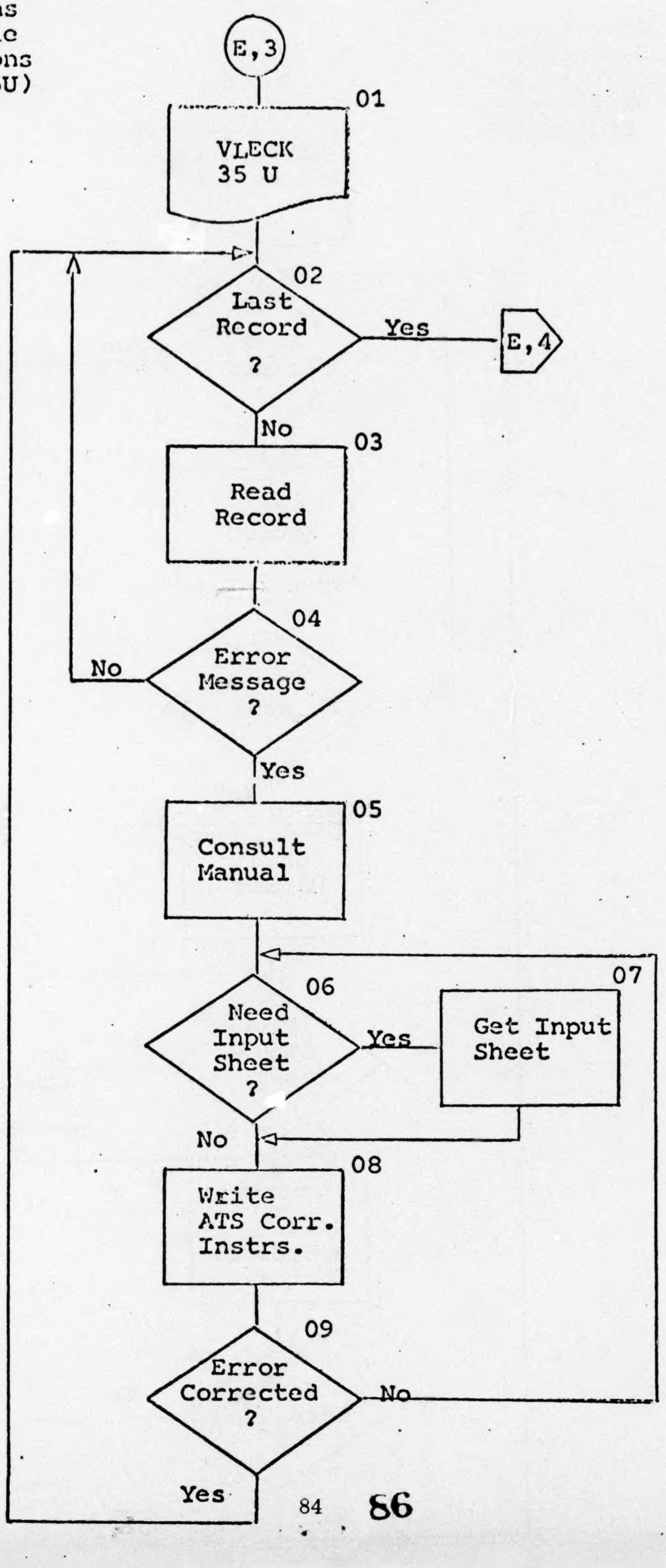
NSPP - Operations
E,1 - Masterfile
Corrections
(ATS Diagnostics)


E,1
01
ATS Diagnostics
02
Last Record ?
Yes
E,2
No
03
Read Record
04
Error Message ?
No
Yes
05
Consult Manual
06
Need Input Sheet ?
Yes
07
Get Input Sheet
No
08
Write ATS Corr. Instrs.
09
Error Corrected ?
No
Yes

NSPP - Operations
E,2 - Masterfile Corrections (Maintenance Diagnostics)


E,2
01
Maintenance Diagnostics
02
Last Record ?
Yes
E,3
No
03
Read Record
04
Error Message ?
No
Yes
05
Consult Manual
06
Need Input Sheet ?
Yes
07
Get Input Sheet
No
08
Write ATS Corr. Instrs.
09
Error Corrected ?
No
Yes

NSPP - Operations
E,3 - Masterfile
Corrections
(VLECK 35U)


E,3
01
VLECK 35 U
02
Last Record ?
Yes
E,4
No
03
Read Record
04
Error Message ?
No
Yes
05
Consult Manual
06
Need Input Sheet ?
Yes
07
Get Input Sheet
No
08
Write ATS Corr. Instrs.
09
Error Corrected ?
No
Yes

NSPP - Operations
E,4 - Masterfile
Corrections
(MED)

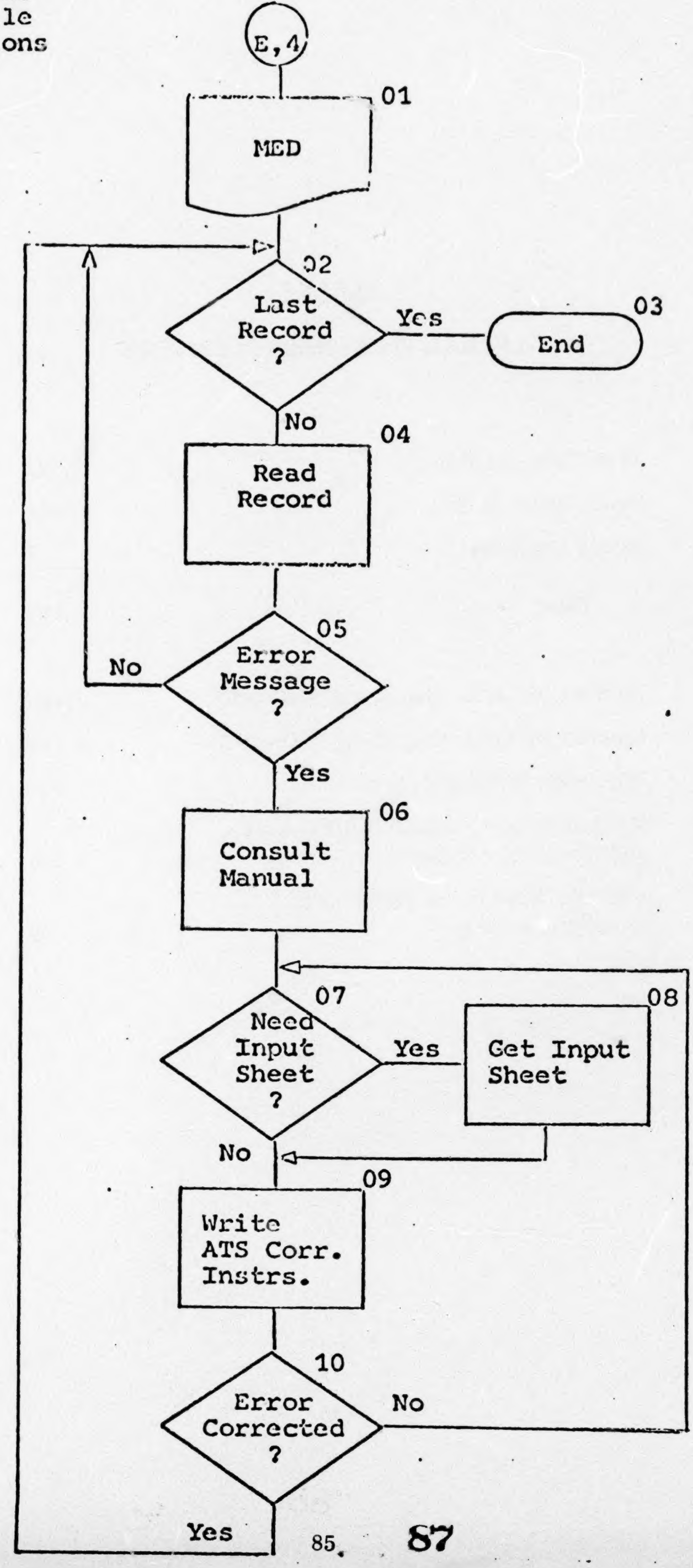

E,4
01
MED
02
Last Record ?
Yes
03
End
No
04
Read Record
05
Error Message ?
No
Yes
06
Consult Manual
07
Need Input Sheet ?
Yes
08
Get Input Sheet
No
09
Write ATS Corr. Instrs.
10
Error Corrected ?
No
Yes

TABLE I

GENERAL FILE CHARACTERISTICS

Live Titles in File	6,221
Dead Titles in File	821
Status Unknown	7
Total	7,049
Number of Titles Beginning 1946-1970	3,866
Number of Titles Beginning 1960-1970	1,746
Publisher is Main Entry	856
Mean Average Number of Characters per Record (rounded)	420
Average Number of Fields per Record (rounded)	9

TABLE II

FREQUENCY DISTRIBUTION OF VARIABLE FIELDS IN FILE

Tag	Field Name	Frequency	Percent of Records	Percent of Tags	Rank
035	Local System Number	7,049	100.00	13.65	1
041	Languages	994	14.10	1.92	12
049	Magnetic Tape Input Source	6,565	93.13	12.71	4
050	LC Call Number	2,840	40.28	5.50	7
060	NLM Call Number	1,046	14.83	2.02	11
070	NAL Call Number	970	13.76	1.87	13
100	Main Entry, Personal	8	0.11	0.01	36
110	Main Entry, Corporate	2,373	33.66	4.59	9
111	Main Entry, Conference	77	1.09	0.14	26
200	Title as it Appears on Piece	4	0.05	0.00	37
240	Uniform Title	--	--	--	--
245	Full Title	7,049	100.00	13.65	2
246	Varying Forms of Title	743	10.54	1.43	16
247	Former Titles and Variants	395	5.60	0.76	18
250	Edition Statement	13	0.18	0.02	35
260	Imprint	6,856	97.26	13.28	3
310	Frequency	149	2.11	0.28	25
321	Former Frequency	69	0.97	0.13	28
362	Dates and Volume Designations	2,384	33.82	4.61	8
500	Note, General	911	12.92	1.76	14

TABLE II (Continued)

Tag	Field Name	Frequency	Percent of Records	Percent of Tags	Rank
510	Note, Indexing and Abstracting	--	--	--	--
515	Note, Numbering Peculiarities	503	7.13	0.97	17
520	Note, Abstract or Annotation	1	0.01	0.00	38
525	Note, Supplement	173	2.45	0.33	24
530	Note, Additional Physical Forms	13	0.18	0.02	34
555	Note, Cumulative Indexes	351	4.97	0.68	19
700	Added Entry, Personal	233	3.30	0.45	22
710	Added Entry, Corporate	3,132	44.43	6.06	5
711	Added Entry, Conference	20	0.28	0.03	32
730	Added Entry, Uniform Title	--	--	--	--
760	Main Series Entry	52	0.73	0.10	30
762	Subseries Entry	17	0.24	0.03	33
765	Original Language Entry	56	0.79	0.10	29
767	Translation Entry	70	0.99	0.13	27
770	Supplement Entry	237	3.36	0.45	21
772	Parent Record Entry	260	3.68	0.50	20
775	Other Editions Available	27	0.38	0.05	31
776	Additional Forms Available	1	0.01	0.00	39
777	"Issued With" Entry	199	2.82	0.38	23
780	Preceding Entry	2,006	28.45	3.88	10
785	Succeeding Entry	844	11.97	1.63	15
850	Library Codes and Holdings	2,932	41.59	5.68	6

TABLE III

LANGUAGES OF PUBLICATION

Afrikaans	1
Arabic	2
Azerbaijani	1
Belorussian	2
Bulgarian	14
Catalan	1
Chinese	24
Croatian	13
Czech	53
Danish	35
Dutch	43
English	3,889
Faroese	1
Finnish	16
Flemish	2
French	498
German	593
Greek	6
Hebrew	1
Hindi	1
Hungarian	25
Indonesian	3
Italian	293
Japanese	100
Kirghiz	1
Latvian	3
Macedonian	2
Moldavian	1
Multilingual	551
Norwegian	15
Polish	80
Portuguese	103
Rumanian	66
Russian	200
Serbian	24
Slovak	9
Spanish	312
Swedish	29
Turkish	9
Ukrainian	21
Uzbek	1

TABLE IV

COUNTRIES OF PUBLICATION

Afghanistan	2
Algeria	10
Angola	2
Argentina	84
Australia	212
Austria	42
Barbados	1
Belgium	89
Bermuda	1
Brazil	89
British Honduras	1
Bulgaria	24
Canada	
Alberta	16
British Columbia	43
Manitoba	6
New Brunswick	1
Newfoundland	1
Nova Scotia	4
Ontario	106
Quebec	33
Saskatchewan	5
Ceylon	2
Chile	16
China, Mainland	20
China, Republic of	19
Colombia	15
Congo (Kinshasa)	2
Costa Rica	1
Cuba	10
Czechoslovakia	86
Denmark	67
Dominican Republic	1
Ecuador	4
El Salvador	1
Ethiopia	1
Faeroe Islands	1
Fiji	2
Finland	35
France	371
Germany, East	112
Germany, West	447
Ghana	3
Greece	8
Guatemala	3
Guinea	1
Guyana	1
Haiti	1
Hong Kong	3
Hungary	51
Iceland	1
India	166
Indonesia	8
Iran	1
Iraq	3
Ireland	11
Israel	12
Italy	317
Ivory Coast	1
Jamaica	5
Japan	195
Kenya	2
Korea, South	2
Lebanon	7
Libya	1
Malagasy Republic	3
Malawi	3
Malaysia	4
Mexico	33
Morocco	3
Mozambique	4
Nepal	1
Netherlands	115
New Caledonia	1
New Zealand	19
Nigeria	11
No Place	192
Norway	27

TABLE IV (Continued)

Pakistan	24	United Kingdom	
Panama	1	England	555
Paraguay	2	Northern Ireland	4
Peru	10	Scotland	26
Philippines	6	Wales	1
Poland	126	United States	
Portugal	25	Alabama	13
Puerto Rico	3	Alaska	23
Rumania	75	Arizona	24
Senegal	1	Arkansas	7
Sierra Leone	1	California	74
Singapore	4	Colorado	15
South Africa	27	Connecticut	23
Southern Rhodesia	5	Delaware	4
Spain	97	District of Columbia	286
Sudan	3	Florida	12
Sweden	82	Georgia	8
Switzerland	160	Hawaii	10
Syria	2	Idaho	1
Tanzania	1	Illinois	210
Thailand	4	Indiana	14
Trinidad and Tobago	1	Iowa	5
Tunisia	3	Kansas	10
Turkey	14	Kentucky	7
Uganda	2	Louisiana	8
U.S.S.R.		Maine	2
Armenian SSR	4	Maryland	100
Azerbaijan SSR	14	Massachusetts	75
Belorussian SSR	6	Michigan	55
Estonian SSR	2	Minnesota	27
Georgian SSR	5	Mississippi	4
Kirghiz SSR	2	Missouri	53
Kazakh SSR	17	Montana	3
Latvian SSR	10	Nebraska	3
Russian SFSR	157	Nevada	3
Tadzhik SSR	2	New Hampshire	5
Turkmen SSR	2	New Jersey	30
Ukrainian SSR	32	New Mexico	6
Uzbek SSR	9	New York	699
United Arab Republic	12	North Carolina	14

TABLE IV (Continued)

United States (Continued)	
North Dakota	1
Ohio	35
Oklahoma	4
Oregon	6
Pennsylvania	179
Rhode Island	3
South Carolina	5
South Dakota	2
Tennessee	12
Texas	21
Utah	8
Vermont	4
Virginia	13
Washington	14
Wisconsin	41
West Virginia	4
Wyoming	1
Uruguay	11
Various Places	3
Venezuela	25
Vietnam, South	2
Yugoslavia	66
Zambia	4